INSTITUTE OF PHYSICS PUBLISHING NONLINEARITY

Nonlinearity **14** (2001) 1071–1104 PII: S0951-7715(01)16985-9

Exact bounds for the polynomial decay of correlation, $1/f$ noise and the CLT for the equilibrium state of a non-Hölder potential

Albert M Fisher[1] **and Artur Lopes**[2]

[1] Dept Mat IME-USP, Caixa Postal 66281, CEP 05315-970 São Paulo, Brazil
[2] Instituto de Matemática, UFRGS, Av. Bento Gonçalves 9500, CEP 91500 Porto Alegre, Brazil

E-mail: afisher@ime.usp.br and alopes@mat.ufrgs.br

Received 9 September 2000, in final form 18 May 2001
Published 24 July 2001
Online at stacks.iop.org/Non/14/1071

Recommended by M Viana

Abstract

We analyse the correlation and limit behaviour of partial sums for the stationary stochastic process $(f(T^t(x)), \mu)$, $t = 0, 1, \ldots$, for functions f of superpolynomial variation, the class $\mathcal{SP}$ defined below (which includes the Hölder functions), where $T : \Sigma^+ \to \Sigma^+$ is the left shift map on $\Sigma^+ = \Pi_0^\infty\{0, 1\}$ and μ is the non-atomic equilibrium measure of a non-Hölder potential $g = g_\gamma$ belonging to a one-parameter family, indexed by $\gamma > 2$.

First, using the renewal equation, we show a polynomial rate of convergence for the associated Ruelle operator for cylinder set observables.

We then use these estimates to prove the following theorems:

- We extend the polynomial convergence for the Ruelle operator to functions $f \in \mathcal{SP}$.
- We show that the measure is weakly Bernouilli and the bounds are polynomial.
- We calculate the decay of correlation of the stationary stochastic process described above, for $f \in \mathcal{SP}$. This decay is polynomial with t: we show in theorem 4.1 an upper bound of the order of $Ct^{2-\gamma}$ when $\gamma > 2$; this estimate is sharp in the sense that for each γ there exist functions f (in fact $f = I_{[0]}$ gives an example) for which one has the lower bound of $ct^{2-\gamma}$ for the decay of its autocorrelation (see theorem 2.8). For the lower bound we use Tauberian theorems. For this example the coefficients decay monotonically, which is important for proving the lower bound.
- Again using Tauberian theorems together with the upper–lower bounds we show that for each $3 > \gamma > 2$ one has the phenomenon of $1/f$ noise for the spectral density of the function $I_{[0]}$ (see theorem 2.8).

- We prove the central limit theorem (CLT) and functional CLT for the case where f is in $\mathcal{SP}$ and for $\gamma > 3$ (theorem 6.5). For this we apply Gordin's method in the setting of a polynomial rate of convergence.

From the perspective of differentiable dynamical systems μ is the unique invariant measure which is absolutely continuous with respect to the Lebesgue measure for an associated doubling map of the circle with an indifferent fixed point. This map $T_1 = T_{1,\gamma}$ is a piecewise linear version of the Manneville–Pomeau map, and the potential g_γ is equal to $-\log DT_{1,\gamma}$.

We emphasize that our class $\mathcal{SP}$ is larger than the classes studied elsewhere.

Mathematics Subject Classification: 37E05

1. Introduction

This work is related to that in 'Invariance principles in log density and convergence to stable flows' (see [FLa]) and 'Self-similar return sets for some maps with an indifferent fixed point' (see [FLb]). We refer the reader to those companion papers for general background, exposition and motivation, and for further references.

We analyse the correlation and limit behaviour of partial sums for the stationary stochastic process $(f(T^t(x)), \mu), t = 0, 1, \ldots,$ for functions f of superpolynomial variation, the class $\mathcal{SP}$ defined at the start of section 2 (which includes the Hölder functions), where $T : \Sigma^+ \to \Sigma^+$ is the left shift map on $\Sigma^+ = \Pi_0^\infty\{0, 1\}$ and μ is the non-atomic equilibrium measure [Wal75, Lop93] of a non-Hölder potential $g = g_\gamma$ belonging to a one-parameter family, indexed by $\gamma > 2$ and described below. These potentials are also not of summable variation (cf [Pol]).

Using renewal theory we calculate the decay of correlation of the stationary stochastic process described above. This decay is polynomial with t: we show in theorem 4.1 for $f \in \mathcal{SP}$ an upper bound of the order of $C_1 t^{2-\gamma}$ in the case $\gamma > 2$; this estimate is sharp in the sense that for each γ there exist functions f (in fact $I_{[0]}$ is an example) for which we show the lower bound of $\overline{C}_1 t^{2-\gamma}$ (see theorem 2.8). In proving the lower bound we make use of Tauberian theorems.

This generalizes some results in [Lop93] (see the appendix of [Lop90]).

Note that, in principle, having a polynomial upper bound does not preclude that the decay is exponential. Indeed, for the expanding map T one can find a Hölder function f such that the stochastic process $X_t = f(T^t(x))$ is independent for the measure μ and therefore the correlation is zero for all t. This is the reason for considering in the exact bound 'some' f and not 'any' f.

For $\gamma > 3$ we prove the central (CLT) limit theorem for functions f in $\mathcal{SP}$ with respect to these equilibrium measures (theorem 6.5). We use a Gordin–Liverani-type argument (see [Via97, Bra88, Liv95]).

The potential g_γ can be viewed as the log-derivative $-\log DT_1$ of a transformation $T_1 = T_{1,\gamma}$ which is a piecewise linear version of a one-dimensional smooth doubling map $T_2 = T_{2,\gamma}$ for $\gamma > 2$ with an indifferent fixed point [Lop90, Lop93, CF90, CG93, CGS92, Bro94, Bro96, Man80]. For further explanation see [FLa]. In figure 1 we show the graphs of the transformations T_1 and T_2. The map T_2 depicted in figure 1 is the Manneville–Pomeau map $T_{2,\gamma}\colon x \mapsto x + x^{1+s} \pmod 1$ where $s = 1/(\gamma - 1)$.

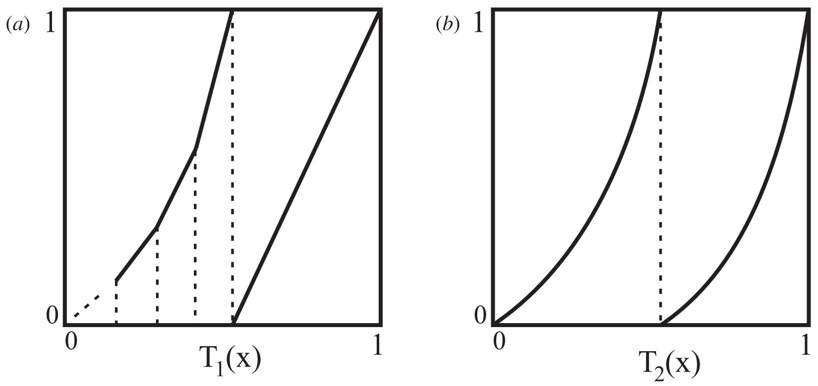

Figure 1. T_1 is the linear by part version of T_2.

Some papers which have results related to those presented here are [BFG99a, BFG99b, Che95, You99, KMS97, Pol, PY, PS92, Mor93, Yur97, Bro94, Bro96, CF90, BG95, FS79, FS88, CG93, CGS92, Pop79, LSV93, LSV99, Lop90, Lop93, LSV99, Iso99, Aar97, ADF92, AF, ADU93].

We mention especially the nice paper [You99] where exact bounds for the convergence to equilibrium (for observables and the Gibbs measure) for maps of Manneville–Pomeau type are proved. Here in order to prove the exact bounds for the correlation coefficients we need a more delicate estimate, see theorem 2.8. We note that this part of the proof makes use of some form of monotonicity.

We refer the reader to [Hof77, Lop90, Lop93] for general properties of the potentials like those studied here. Hofbauer's work was part of the original inspiration for this paper.

In section 6 we show the central limit theorem (with Gaussian limiting distribution; theorem 6.5), for $\gamma > 3$. The method breaks down for $\gamma \leqslant 3$.

There is a good reason for this: for the cases $2 < \gamma < 3$ and $1 < \gamma < 2$ one obtains as limiting distributions the completely asymmetric stable laws of index $\alpha = \gamma - 1$ and the Mittag–Leffler laws, respectively (see [FLa, FLb] and also [Fel49]). In fact, for the transition value $\gamma = 3$ the central limit theorem also holds (see [FLa]); however, we are not able to prove this using the techniques of the present paper.

For these potentials (for $\gamma > 2$), unlike the classical (Hölder) case, there is no unique equilibrium state: there are two such probability measures, the measure we are interested in, denoted μ, and point mass at the point $(111\ldots) \in \Sigma^+$ [Hof77, Lop93].

From our point of view, for $\gamma \in (1, 2)$, the natural measure is infinite and sigma-finite; it is the unique invariant absolutely continuous measure (up to multiplication by a constant), however, it no longer fits the usual definition of an equilibrium state (of which only one now exists, the point mass [Hof77]). For this natural measure, logarithmic averages are known to exist [ADF92], as well as invariance principles in log density (for all $\gamma > 0$) (see [FLa]); in this paper we restrict attention to the finite measure case and hence to $\gamma > 2$.

A key tool we use in obtaining our principal estimates is renewal theory [KT75]; in one part of the proof we also need a more delicate control via Tauberian theorems (see the comments before theorem 2.8).

An application of the polynomial decay of correlation is the following: a classical result in trigonometric series relates the rate of decay to the zero of the Fourier coefficients a_n of a function $F(\lambda) = \sum_{n\geqslant 0} a_n \cos(n\lambda)$, $\lambda \in [-\pi, \pi]$ with the property of F being Hölder or at least in class L^p as a function of the variable λ. In [Bar], section 10 and [Pos79] this question is studied.

Consider a function h with μ-mean zero and $a_n, n \geqslant 0$ the correlation coefficients of the stationary stochastic process $X_n = h(T^n)$ with respect to μ, that is $a_n = \int h(T^n(x))h(x)\,\mathrm{d}\mu(x)$. In this case the function $F(\lambda) = \sum_{n\geqslant 0} a_n \cos(n\lambda)$ is called the spectral density function associated with h and μ (see, for example, [KT75]). From the precise estimates obtained for our example and Tauberian theorems we show that for each $3 > \gamma > 2$ then for the function $h(x) = I_{[0]}(x) - \mu[0]$, the spectral density $F(\lambda)$ has the '$\frac{1}{f}$-noise' property. Precisely, we show that $F(\lambda)$ is of the order of $\lambda^{-\beta}, \beta = 3 - \gamma > 0$, for λ approaching zero. For full details see theorem 2.8. The terminology $1/f$ used in the literature is perhaps a bit confusing; f there denotes frequency, which in our notation is denoted instead by λ; in that other notation, we are, in fact, showing the '$1/f^{3-\gamma}$-noise' property.

2. Polynomial convergence of the Ruelle operator

Shift space; superpolynomial variation

We introduce a new class of functions more general than the Hölder functions.

We define $\Sigma^+ = \Pi_0^\infty\{0, 1\}$ and denote the left shift map on Σ^+ by T. We write $z = (z_0 z_1 \ldots)$ for a point in Σ^+ and define $[w_0 w_1 \ldots w_k] = \{z\colon z_0 = w_0, z_1 = w_1, \ldots, z_k = w_k\}$; this is called a *k-cylinder set*. The collection of all k-cylinder sets will be denoted by $\mathcal{C}_k$. We give the space Σ^+ the product topology, and define the (compatible) metric $d(x, y) = 2^{-n}$, where n is the greatest integer such that x, y lie in the same n-cylinder set. The Borel σ-algebra of Σ^+ is denoted by $\mathcal{B}$. We write $\mathcal{C}(\Sigma^+)$ for the set of continuous real-valued functions on Σ^+. We define the kth *variation* of $f \in \mathcal{C}(\Sigma^+)$,

$$\mathrm{var}_k(f) = \sup\{|f(x) - f(y)|\colon x, y \in C \in \mathcal{C}_k\}.$$

For $\alpha \in (0, 1)$, we define the class of functions whose variation is exponentially small as a function of k, for base α:

$$\mathcal{H}_\alpha \equiv \{f\colon \exists c > 0 \text{ with } \mathrm{var}_k f \leqslant c\alpha^k\},\ \text{ for all } k \geqslant 0\}.$$

These are the Hölder functions with exponent $-\log\alpha/\log 2$ with respect to the metric d. We write $\|\cdot\|_\infty$ for the sup norm on $\mathcal{C}(\Sigma^+)$, and define a norm on $\mathcal{H}_\alpha \subseteq \mathcal{C}(\Sigma^+)$ by

$$\|f\|_\alpha \equiv \|f\|_\infty + c$$

where c is the inf of the possible Hölder constants for that exponent (or equivalently, for that base α); thus,

$$c = \sup_k\{\mathrm{var}_k f \cdot \alpha^{-k}\}.$$

In the standard setting [Bow75], Hölder functions appear in two ways: as potential functions, which are then used to define a Ruelle operator and from that the relevant invariant measure on the space (the *Gibbs* or *equilibrium state*), and as an observable, making a measurement on the system and for which one wants to study time averages, correlations and so on. In our case, the potential will come from a specific one-parameter family of non-Hölder potentials $g = g_\gamma$ for $\gamma > 2$. Our allowed observables will come from a class $\mathcal{SP}$, the functions with superpolynomial variation, defined as follows. First, for $a > 0$, we define

$\underline{c_a} = \underline{c_a}(f) = \limsup_k (\mathrm{var}_k\, f/k^{-a})$ and write for the functions with polynomial variation of exponent a

$$\mathcal{P}_{(a)} \equiv \{f\colon \underline{c_a}(f) < \infty\}.$$

Thus for $f \in \mathcal{P}_{(a)}$ we have that for every $\delta > 0$, $\mathrm{var}_k\, f \leqslant \underline{c_a}(1+\delta)k^{-a}$, for all k larger than a number which depends on δ and f. For the choice $\delta = 1$, we will write $\underline{k_a} = \underline{k_a}(f)$ for this number.

We define a norm on $\mathcal{P}_{(a)}$ by

$$\|f\|_{(a)} \equiv \|f\|_\infty + \underline{c_a}.$$

Next, we write

$$\mathcal{SP} = \{f \in \mathcal{C}(\Sigma^+)\colon \sup_{a>0}\{\|f\|_{(a)}\} < \infty\}.$$

This is larger than the collection of all Hölder functions; one has

$$\cup_{\alpha\in(0,1)}\mathcal{H}_\alpha \subsetneq \mathcal{SP} \subsetneq \cap_{a>0}\mathcal{P}_{(a)}.$$

We define on $\mathcal{SP}$

$$\|f\|_{\mathcal{SP}} \equiv \sup_{a>0}\{\|f\|_{(a)}\}$$

which is clearly a norm.

Note that we have for $f \in \mathcal{SP}$, for each $a > 0$,

$$\mathrm{var}_k\, f \leqslant 2\|f\|_{\mathcal{SP}} k^{-a} \tag{1}$$

for all $k > \underline{k_a}(f)$.

Invariant measures and the Ruelle operator

Notation. We write $a_n \approx b_n$ if $\lim_{n\to\infty} a_n/b_n = 1$.

We will also use the following notation: the statement, $a_n < C_2 n^{\delta^+}$ means '*for* $\delta + \varepsilon$*, for every* $\varepsilon > 0$', and $a_n > C_2 n^{\delta^-}$ means '*for* $\delta - \varepsilon$*, for every* $\varepsilon > 0$'.

We write I_A for the indicator function of a set A; thus $I_A(z) = 1$ if $z \in A$ and $= 0$ if $z \notin A$.

We denote by $M_k \subset \Sigma^+$, for $k \geqslant 1$, the cylinder set $[\underbrace{111\ldots11}_{k}0]$ and by M_0 the cylinder set $[0]$. The ordered collection $(M_k)_{k=0}^\infty$ is a partition of Σ^+; in other words these sets are disjoint and their union is the whole space (minus the point $(111\ldots)$). Note that T maps M_k bijectively onto M_{k-1} for $k \geqslant 1$, and onto Σ^+ for $k = 0$. This partition therefore gives a second nice symbolic dynamics for the map T, as a countable state Markov chain (see [FLa]).

For $\gamma > 1$ a fixed constant, we consider the potential $g(x)$ such that $g(111\ldots) = 0$,

$$g(x) = \begin{cases} a_k = -\gamma \log\left(\frac{k+1}{k}\right) & \text{for } x \in M_k \quad \text{with } k \neq 0 \\ a_0 = -\log(\zeta(\gamma)) & \text{for } x \in M_0 \end{cases}$$

where ζ is the Riemann zeta function. Note that g is a continuous function on Σ^+, since $g(x) \to 0$ as $x \to (111\ldots)$ in Σ^+.

We observe that $0 < \mathrm{e}^{g(x)} < 1$ for all $x \in \Sigma^+$ except for $(111\ldots)$, where $\mathrm{e}^{g(x)} = 1$.

By definition, $\zeta(\gamma) = (1^{-\gamma} + 2^{-\gamma} + \cdots)$ and so the reason for the above apparently mysterious value of a_0 becomes clear: this choice just guarantees that the a_k are normalized in the following sense. Defining $s_k = a_0 + a_1 + \cdots + a_k$, note that

$$\mathrm{e}^{s_k} = \mathrm{e}^{a_0}(k+1)^{-\gamma} = \frac{(k+1)^{-\gamma}}{\zeta(\gamma)} \tag{2}$$

hence $\sum_{k=0}^{\infty} \mathrm{e}^{s_k} = 1$. This normalization has a geometrical meaning: it implies that the potential can be derived in a natural way from a piecewise linear map of the interval (see [FLa]).

These potentials are not Hölder and, in fact, are not of summable variation. It is known [Hof77] that the pressure $P(g) = 0$ and that there exist two equilibrium states for such a potential g: point mass (the Dirac delta) at $(111\ldots)$, and a second measure which we shall denote by μ. Note that in [FLa] the roles of 0 and 1 have been reversed, to make the correspondence with Manneville–Pomeau maps clearer; here we are sticking closer to the usage of [Hof77].

We write partial sums as

$$S_0\phi(x) = 0,\ S_1\phi(x) = \phi(x), \ldots, S_n\phi(x) = \phi(x) + \phi(T(x)) + \cdots + \phi(T^{n-1}(x)).$$

We recall that the Ruelle operator $\mathcal{L}_\phi$ is defined by $\mathcal{L}_\phi^k f(z) = \sum_{y:\, T^k(y)=z} \mathrm{e}^{S_k\phi(y)} f(y)$. The dual operator $\mathcal{L}_\phi^*$ acts on a measure m by $(\mathcal{L}_\phi^* m)(f) = m(\mathcal{L}_\phi f) = \int \mathcal{L}_\phi f \,\mathrm{d}m$. One knows that the following are equivalent for a measure μ and a measurable function ϕ on Σ^+ (see [Kea72, Led74, PP90]).

(a) μ is invariant;
(b) the local Radon–Nikodym derivative $(\mathrm{d}\mu \circ \sigma)/\mathrm{d}\mu = \mathrm{e}^{-\phi} = 1/p$ is normalized, i.e. $\sum_{y:\, T^k(y)=z} p(y) = 1$ (this is a g-measure in the terminology of Keane, for $g = p$);
(c) for a normalized potential ϕ, $\mathcal{L}_\phi^*(\mu) = \mu$.

We shall need the following:

Lemma 2.1. *Let ϕ be a continuous non-zero potential which is constant on the sets $M_0, M_1, \ldots$. Assume m is a probability measure such that $\mathcal{L}_\phi^* m = m$. Then for any cylinder set $[w_0 \ldots w_k] = [w_0 \ldots w_{k-1}0]$,*

$$m[w_0 \ldots w_{k-1}0] = \mathrm{e}^{S_k\phi(w_0\ldots w_{k-1}0)} m[0] \tag{i}$$

and for all $t \geqslant k$,

$$\mathcal{L}_\phi^t I_{[w_0\ldots w_{k-1}0]} = \mathrm{e}^{S_k\phi(w_0\ldots w_{k-1}0)} \mathcal{L}_\phi^{t-k} I_{[0]}. \tag{ii}$$

Proof. Note that any set of the form $[w_0 \ldots w_{k-1}0]$ is completely contained in some M_i, hence by the assumption ϕ is constant there. Now for $y = (y_0 y_1 \ldots)$, $T^k(y) = z$ iff $(y_k y_{k+1} \ldots) = (z_0 z_1 \ldots)$, so therefore $I_{[w_0\ldots w_k]}(y) = 1$ and $T^k(y) = z$ iff $y = (w_0 \ldots w_k z_1 \ldots)$ and $z_0 = w_k$. Hence

$$\begin{aligned} \mathcal{L}_\phi^k I_{[w_0\ldots w_k]}(z) &\equiv \sum_{y:\, T^k(y)=z} \mathrm{e}^{S_k\phi(y)} I_{[w_0\ldots w_k]}(y) = \mathrm{e}^{S_k\phi(w_0\ldots w_k z_1\ldots)} I_{[w_k]}(z) \\ &= \mathrm{e}^{S_k\phi(w_0\ldots w_{k-1}0)} I_{[w_k]}(z) = \mathrm{e}^{S_k\phi(w_0\ldots w_{k-1}0)} I_{[0]}(z). \end{aligned}$$

Therefore, $m[w_0 \ldots w_k] = (\mathcal{L}_\phi^{*k} m)[w_0 \ldots w_k] = m(\mathcal{L}_\phi^k I_{[w_0\ldots w_k]}) = \mathrm{e}^{S_k\phi(w_0\ldots w_{k-1}0)} m[0]$, giving statement (a).

The above includes a proof of (b) for the case $t = k$; for $t > k$ the result then follows, since $\mathcal{L}_\phi^t = \mathcal{L}_\phi^{t-k} \circ \mathcal{L}_\phi^k$. $\square$

Theorem 2.2 (Perron–Frobenius theorem for the potentials g_γ). *Assume $\gamma > 2$. There exists a unique probability measure ν on Σ^+ which satisfies $\mathcal{L}_g^*\nu = \nu$.*

There is a unique continuous extended-real-valued function h with $\mathcal{L}_g h = h$, normalized so $\nu(h) = 1$; h is finite except at $(111\ldots)$ where it has value $+\infty$.

Proof. First we prove uniqueness of a measure satisfying the above conditions. As a first step we show that such a ν can have no atomic part, i.e. can contain no point masses. We write $\nu = \nu_a + \nu_c$ for the atomic and continuous parts; we want to show $\nu_a = 0$.

$\mathcal{L}_g^*$ acts on point masses as

$$\mathcal{L}_g^*(\delta_x) = \mathrm{e}^{g(y)}\delta_y + \mathrm{e}^{g(w)}\delta_w \tag{3}$$

where $T^{-1}(x) = \{y, w\}$. Now $\nu = \mathcal{L}_g^*(\nu) = \mathcal{L}_g^*(\nu_a + \nu_c) = \mathcal{L}_g^*(\nu_a) + \mathcal{L}_g^*(\nu_c)$, hence $\mathcal{L}_g^*$ preserves ν_a and ν_c. If ν_a has some mass at the point $w = (w_0 w_1 \ldots)$, it also has mass at all points in the *grand orbit* of w, all the pre- and forward images of w, by 1. Now from lemma 2.1 since $\mathrm{e}^{g(z)} < 1$ for all $z \neq (111\ldots)$ in Σ^+, where the value is 1, the measures of cylinders $\nu[w_0 \ldots w_{k-1}0] \to 0$ as $k \to \infty$, unless w_i is eventually $111\ldots$, i.e. unless w is in the grand orbit of $(111\ldots)$. Therefore, if $\nu_a \neq 0$, it has positive mass at that fixed point. We have

$$\begin{aligned}\nu_a(111\ldots) = \mathcal{L}_g^*\nu_a(111\ldots) &= \mathrm{e}^{g(0111\ldots)}\nu_a(0111\ldots) + \mathrm{e}^{g(111\ldots)}\nu_a(111\ldots)\\ &= \mathrm{e}^{g(0111\ldots)}\nu_a(0111\ldots) + \nu_a(111\ldots)\end{aligned}$$

yet we know $\mathrm{e}^{g(0111\ldots)} > 0$ and $\nu_a(0111\ldots) > 0$, giving a contradiction. Hence $\nu_a = 0$ and the measure is purely continuous.

From lemma 2.1, for any cylinder set of the form $[w_0 \ldots w_{k-1}0]$, we have

$$\nu[w_0 \ldots w_{k-1}0] = \mathrm{e}^{S_k g(w_0 \ldots w_{k-1}0)}\nu[0].$$

Hence (since all cylinders are unions of sets of that form) we will know that ν is unique if we can show that $\nu[0]$ is uniquely defined by the conditions. Now in particular the above equation says that, writing for $k \geqslant 0$,

$$\nu_k \equiv \nu(M_k) = \nu[\underbrace{11\ldots1}_{k}0]$$

we have

$$\nu_k = \exp\left\{S_{k+1}g(\underbrace{11\ldots1}_{k-1}0)\right\}\nu[0] = \mathrm{e}^{s_k - a_0}\nu_0.$$

By assumption, ν is a probability measure; we have equivalently (since, as we now know, there is no mass at $(111\ldots)$), $\sum_{k=0}^{\infty}\nu_k = 1$. Hence

$$1 = \left(\sum_{k=1}^{\infty}\mathrm{e}^{s_k - a_0}\right)\nu_0 + \nu_0 = \nu_0\left(\sum_{k=0}^{\infty}\mathrm{e}^{s_k}\right)\mathrm{e}^{-a_0}.$$

Therefore,

$$\nu_0 = \nu(M_0) = \mathrm{e}^{s_0} = \mathrm{e}^{a_0} = \frac{1}{\zeta(\gamma)}.$$

Hence ν is determined and the eigenmeasure is unique. Note that the above formula therefore tells us that, using (2),

$$\nu_k = \nu(M_k) = \mathrm{e}^{s_k} = \frac{(k+1)^{-\gamma}}{\zeta(\gamma)} \qquad \text{for} \quad k \geqslant 0. \tag{4}$$

Next we prove existence. We define ν as above on cylinder sets of the form $[w_0 \dots w_{k-1}0]$; for sets ending in a 1, i.e. of the form $[w_0 \dots w_{k-1}1]$, we note that

$$[w_0 \dots w_{k-1}1] = [w_0 \dots w_{k-1}10] \cup [w_0 \dots w_{k-1}110] \cup \cdots$$

and define the measure to be the sum of these. To prove existence it is sufficient to show that this definition is additive on k-cylinder sets, for each k. In fact, the argument just given contains the proof that $\nu[0] + \nu[1] \equiv \nu[0] + \nu[10] + \nu[110] + \cdots = 1$, so ν is additive on 1-cylinders. One can show inductively that this calculation applies to the case $k \geqslant 1$; we leave details to the reader (see also [FLa]).

Next we construct an eigenfunction h for $\mathcal{L}_g$, with eigenvalue 1. We will write h_k for the value of h on M_k. From the equation $\mathcal{L}_g h = h$ we know that h is an eigenfunction iff it satisfies this sequence of equations: $e^{a_0}h_0 + e^{a_1}h_1 = h_0$, $e^{a_0}h_0 + e^{a_2}h_2 = h_1$, $e^{a_0}h_0 + e^{a_3}h_3 = h_3, \dots$ and so $h_1 = (h_0 - e^{a_0}h_0)e^{-a_1}$, and for $k \geqslant 1$, $h_k = (h_{k-1} - e^{a_0}h_0)e^{-a_k}$.

We define such a function $\tilde{h}$ by choosing arbitrarily

$$\tilde{h}_0 = (1^{-\gamma} + 2^{-\gamma} + \cdots) = e^{-a_0} \equiv \zeta(\gamma) > 1$$

and then setting

$$\tilde{h}_1 = 2^{\gamma}(2^{-\gamma} + 3^{-\gamma} + \cdots)$$
$$\tilde{h}_2 = 3^{\gamma}(3^{-\gamma} + 4^{-\gamma} + \cdots)$$

and so on. This defines $\tilde{h}$ on $\Sigma^+ \setminus \{(111\dots)\}$; we set $\tilde{h}(111\dots) = \infty$. Then $\tilde{h}$ is an eigenfunction, satisfying $\mathcal{L}_g(\tilde{h}) = \tilde{h}$; what we need to check at the point $(111\dots)$ is that

$$(\mathcal{L}_g\tilde{h})(111\dots) = e^{g(0111\dots)}\tilde{h}(0111\dots) + e^{g(111\dots)}\tilde{h}(111\dots) = \tilde{h}(111\dots).$$

This is indeed valid, since $\tilde{h}(111\dots) = \infty$.

Next we will verify that $\tilde{h}$ is a continuous extended-real-valued function; for this we estimate its values near $(111\dots)$; we will also need these computations later on.

Our formula for $\tilde{h}(x)$, for $x \in M_t$, is

$$\tilde{h}(x) = \tilde{h}_t = \nu_t^{-1} \sum_{i=t}^{\infty} \nu_i. \tag{5}$$

Since (for $\gamma > 1$)

$$\int_{t+1}^{\infty} x^{-\gamma}\,dx \leqslant \sum_{i=t}^{\infty} i^{-\gamma} \leqslant \int_{t}^{\infty} x^{-\gamma}\,dx$$

one has

$$\frac{(t+1)^{1-\gamma}}{\gamma - 1} \leqslant \sum_{i=t}^{\infty} i^{-\gamma} \leqslant \frac{t^{1-\gamma}}{\gamma - 1} \leqslant t^{1-\gamma} \tag{6}$$

with the last inequality holding for $\gamma > 2$.

Since $(t+1)/t \to 1$, we have that

$$\sum_{i=t}^{\infty} i^{-\gamma} \approx \frac{t^{1-\gamma}}{\gamma - 1} \approx \frac{(t+1)^{1-\gamma}}{\gamma - 1}. \tag{7}$$

Therefore,

$$\tilde{h}_t \equiv (t+1)^{\gamma} \sum_{i=t+1}^{\infty} i^{-\gamma} \approx \frac{t+1}{\gamma - 1}. \tag{8}$$

Thus $\tilde{h}(x) \to \infty$ as $x \to (111\ldots)$, so $\tilde{h}$ is continuous, as claimed. Finally, we note that the integral $\int \tilde{h}(x)\,\mathrm{d}\nu(x)$ is finite if and only if $\gamma > 2$ (since $\tilde{h}_n \nu_n \approx (n+1)^{1-\gamma}$). For $\gamma > 2$, one can normalize $\tilde{h}$, multiplying by a constant u to obtain $h = u\tilde{h}$ with $\int h\,\mathrm{d}\nu = 1$.

This completes the proof of the theorem. □

The value of the above constant is

$$u = \frac{1}{\sum_{t=1}^{\infty} t\nu_{t-1}} = \frac{\zeta(\gamma)}{\sum_{t=1}^{\infty} t^{1-\gamma}} = \frac{\zeta(\gamma)}{\zeta(\gamma-1)}. \tag{9}$$

Note that $u\zeta(\gamma) = h_0 = h(x), \forall x \in M_0$.

Consider now the invariant probability measure $\mu = h\nu$. We have, from (4), (5) and (9), that for $t \geqslant 0$,

$$\mu_t = \mu(M_t) = \mu[\underbrace{111\ldots11}_{t}0] = \nu_t\, u\, \tilde{h}_t = u\sum_{i=t}^{\infty} \nu_i = \frac{1}{\zeta(\gamma-1)} \sum_{i=t+1}^{\infty} i^{-\gamma}. \tag{10}$$

Note that

$$\mu(M_0) = \frac{\zeta(\gamma)}{\zeta(\gamma-1)}.$$

From (6) we have therefore the following upper and lower bounds: for all $t \geqslant 1$,

$$\frac{(t+2)^{1-\gamma}}{(\gamma-1)\zeta(\gamma-1)} \leqslant \mu_t \leqslant \frac{(t+1)^{1-\gamma}}{(\gamma-1)\zeta(\gamma-1)} \leqslant \frac{t^{1-\gamma}}{(\gamma-1)\zeta(\gamma-1)}. \tag{11}$$

For $\gamma > 2$ the constant is < 1 so this gives

$$\mu_t \leqslant (t+1)^{1-\gamma} \leqslant t^{1-\gamma}. \tag{12}$$

As $\mathcal{L}_g(h) = h$, then

$$\sum_{T(x)=y} \mathrm{e}^{g(x)} h(x) = h(y)$$

and therefore

$$\sum_{T(x)=y} \mathrm{e}^{g(x)} \frac{h(x)}{h(y)} = \sum_{T(x)=y} \mathrm{e}^{g(x)} \frac{h(x)}{h(T(x))} = 1. \tag{13}$$

We define a second potential ψ by:

$$\psi(x) = g(x) + \log h(x) - \log h(T(x)) \tag{14}$$

where we define $\infty - \infty = 0$, and write

$$p(x) = \mathrm{e}^{\psi(x)} = \mathrm{e}^{g(x)} \frac{h(x)}{h(T(x))} \tag{15}$$

we note that $p(111\ldots) = 1$ and $p(0111\ldots) = 0$.

Replacing the potential g by the cohomologous potential ψ does not change the equilibrium state (Gibbs state). This makes some of our calculations easier, but some also harder, as the potential now depends on two symbols of the process generated by the partition $(M_k)_{k=0}^{\infty}$.

Theorem 2.3. *Let $\gamma > 2$. The measure $\mu = h \cdot \nu$ is an invariant probability measure. It is the unique normalized eigenmeasure for eigenvalue 1 for $\mathcal{L}_{\psi}^{*}$, and is a p-balanced measure for $p = \mathrm{e}^{\psi}$. There is a unique normalized eigenfunction for eigenvalue 1, the constant function 1.*

Proof. A simple computation using lemma 2.1 shows $\mu = h\nu$ is invariant. (For other proofs see lemma 1.13 in [Bow75], or for the present case, [FLa].) $\square$

Theorem 2.4 (See [Hof77]). *There are exactly two equilibrium states for the potential* g_γ, $\gamma > 2$: *the measure* μ *and point mass at* $(111\ldots)$.

Remark 2.1. The point mass at $(111\ldots)$ is T-invariant and does satisfy the definition of an equilibrium state, but it is not invariant for $\mathcal{L}_\psi^*$, as a consequence of theorem 2.2 above.

Note that $h(x)$ and $g(x)$ are constant in each interval M_t and that the value $\psi(x)$ depends on $g(x)$, $h(x)$ and $h(T(x))$.

Lemma 2.5. *For the potential* ψ, *for any cylinder set* $[w_0 \ldots w_{k-1}0]$ *where* $k \geqslant 1$,

$$\mu[w_0 \ldots w_{k-1}0] = \mathrm{e}^{S_k\psi(w_0\ldots w_{k-1}0)}\mu[0] \tag{i}$$

and for all $t \geqslant k$,

$$\mathcal{L}_\psi^t I_{[w_0\ldots w_{k-1}0]} = \mathrm{e}^{S_k\psi(w_0\ldots w_{k-1}0)}\mathcal{L}_\psi^{t-k} I_{[0]}. \tag{ii}$$

Hence

$$\frac{\mu[w_0 \ldots w_{k-1}0]}{\mu[0]} = \mathrm{e}^{S_k\psi(w_0\ldots w_{k-1}0)} \tag{iii}$$

and

$$\mathcal{L}_\psi^t I_{[w_0\ldots w_{k-1}0]} - \mu[w_0 \ldots w_{k-1}0] = \frac{\mu[w_0 \ldots w_{k-1}0]}{\mu[0]}\big(\mathcal{L}_\psi^{t-k} I_{[0]} - \mu([0])\big). \tag{iv}$$

We now calculate explicitly $\psi(x)$, which now depends on which element of $(M_k)_0^\infty$ the point x lies in at time 1 as well as at time 0.

Recall that for $x \in M_n$ where $n \geqslant 1$, we know that $T(x) \in M_{n-1}$. Therefore, $\psi(x)$ is constant on M_n when $n \geqslant 1$; we shall write $\psi_n \equiv \psi(x)$ in this case. For $x \in M_0$ on the other hand, all possible future states $M_0, M_1, \ldots$ can occur. There we write $\psi_0^k = \psi(x)$ for this value, when $x \in M_0 \cap T^{-1}(M_k)$ where $k \neq 0$. We shall also use a second notation for these values: since $[00] = M_0 \cap T^{-1}M_0$, $[010] = M_0 \cap T^{-1}M_1$, $[0110] = M_0 \cap T^{-1}M_2$, and $[10] = M_1$, $[110] = M_2, \ldots$, we write, for example, $\psi(110) = \psi_2 = \psi(110\ldots)$, and $\psi(0110) = \psi_0^2 = \psi(0110\ldots)$.

Assume first that $x \in M_n$ where $n \geqslant 1$. Since

$$\frac{h}{h\circ T} = \frac{\tilde{h}}{\tilde{h}\circ T}$$

we have, using (7):

$$\mathrm{e}^{\psi_n} \equiv \exp\Big\{\psi(\underbrace{11\ldots1}_{n}0)\Big\} \equiv \mathrm{e}^{g_n}\frac{h_n}{h_{n-1}} = \left(\frac{n+1}{n}\right)^{-\gamma}\frac{\tilde{h}_n}{\tilde{h}_{n-1}} \approx \left(\frac{n+1}{n}\right)^{1-\gamma}. \tag{16}$$

For $x \in M_0 \cap T^{-1}(M_k)$ with $k \neq 0$ we have

$$\mathrm{e}^{\psi_0^k} \equiv \exp\Big\{\psi(0\underbrace{11\ldots1}_{k}0)\Big\} \equiv \mathrm{e}^{a_0}\frac{\tilde{h}_0}{\tilde{h}_k} = \frac{1}{\tilde{h}_k} \approx \frac{\gamma-1}{k+1} \tag{17}$$

again using (7).

Lastly, for $x \in M_0 \cap T^{-1}M_0$,

$$e^{\psi(x)} \equiv e^{\psi_0^0} = e^{g_0}\frac{\tilde{h}_0}{\tilde{h}_0} \tag{18}$$

$$= e^{g_0} = e^{a_0} = \frac{1}{\zeta(\gamma)}. \tag{19}$$

Recall that γ^+ by definition means: *for* $\gamma + \varepsilon$, *for every* $\varepsilon > 0$.

Now we come to the main results of this section, where we describe the behaviour of the Ruelle operator, at points in M_s, when applied to the indicator function of a cylinder set.

Estimates for the cylinder set [1]

We will show:

Theorem 2.6.

(a) For $\gamma > 2$, *there exists a constant* $\hat{c}_\gamma$ *such that for all* $z \in M_s$, *for* $s \geqslant 0$,

$$|\mu[1] - \mathcal{L}_\psi^t I_{[1]}(z)| \leqslant \hat{c}_\gamma\big((s+t)^{2-\gamma} + B(s,t)\big) \tag{20}$$

where

$$B(0,t) = 0 \qquad \text{and} \qquad B(s,t) = \left(\frac{s+t}{s}\right)^{1-\gamma} \qquad \text{for} \quad s \geqslant 1.$$

(b) For $z \in M_0$, *we have furthermore that* $\exists \overline{C}_1, C_1 > 0$ *such that*

$$\overline{C}_1 t^{2-\gamma} \leqslant \mu[1] - \mathcal{L}_\psi^t I_{[1]}(z) \leqslant C_1 t^{2-\gamma}. \tag{21}$$

Proof. For $t \geqslant 1$, we have

$$\begin{aligned}(\mathcal{L}_\psi^t I_{[1]})(z) &= \sum_{y_0\cdots y_{t-1}} e^{S_t\psi(y_0\ldots y_{t-1}z)} I_{[1]}(y_0\ldots y_{t-1}z)\\ &= \sum_{y_1\cdots y_{t-1}} e^{S_t\psi(1y_1\ldots y_{t-1}z)}\\ &= \sum_{y_{t-1}=0}^{1} e^{\psi(y_{t-1}z)} \sum_{y_{t-2}=0}^{1} e^{\psi(y_{t-2}y_{t-1}z)} \cdots \sum_{y_1=0}^{1} e^{\psi(y_1\ldots y_{t-1}z)}\big(e^{\psi(1y_1\ldots y_{t-1}z)}\big).\end{aligned} \tag{22}$$

Let t be fixed with $t \geqslant 1$.

We claim that the function $\mathcal{L}_\psi^t I_{[1]}(z)$ is constant on M_0, i.e. that for $z = (z_0z_1z_2\ldots) = (0z_1z_2\ldots)$ the value does not depend on $z_1, z_2, \ldots$. We have previously seen that $\psi(x)$ is constant on any cylinder of the form $[***0]$, where the $*$ denotes 0 or 1 and the number of $*$s is $\geqslant 1$. Consequently, for each quantity in the above expression, if $z_0 = 0$ one need look no further to ascertain the corresponding value; this verifies the claim.

For $z \notin M_0$, the same reasoning applies: now we need look no further than the first 0 which occurs in $z = (z_0z_1z_2\ldots)$. Supposing z is of the form $z = (\underbrace{11\ldots1}_{s}0z_{s+1}\ldots)$, then any part of the above expression of the form $e^{\psi(y_l\ldots y_{t-1}z)}$ does not depend on $z_{s+1}, z_{s+2}, \ldots$ (Note that if a 0 already occurs in the y_i, then we can of course stop earlier.) The points z of that form are exactly those in M_s for $s \geqslant 1$.

In summary, for t fixed, and $t \geqslant 1$, $\mathcal{L}_\psi^t I_{[1]}(z)$ is constant on each M_n for $n \geqslant 0$.

We write this value as follows, where $z \in M_s$ for $s \geqslant 0$:

$$A_t^s = \mathcal{L}_\psi^t I_{[1]}(z). \tag{23}$$

For $s = 0$ we shall also write $A_t \equiv A_t^0$. □

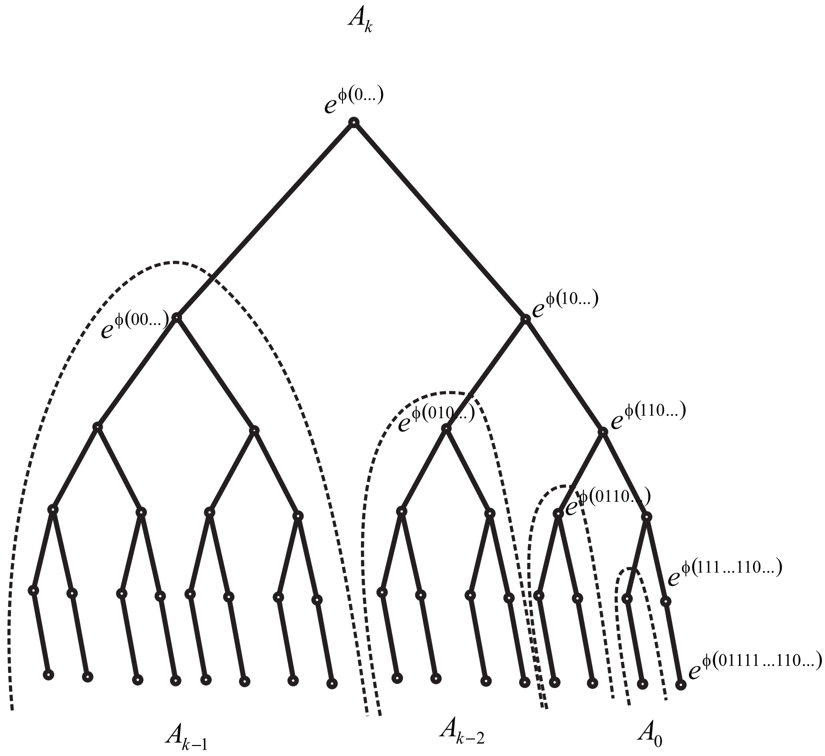

Figure 2. Geometrical view of the renewal equation.

The case $z \in M_0$

From (22) and figure 2 one can see that $A_t \equiv A_t^0$ with $t \geqslant 1$ satisfies

$$A_t = A_{t-1}(e^{\psi_0^0}) + A_{t-2}(e^{\psi_0^1}e^{\psi_1}) + A_{t-3}(e^{\psi_0^2}e^{\psi_2}e^{\psi_1}) + \cdots$$
$$+A_1(e^{\psi_0^{t-2}}e^{\psi_{t-3}}e^{\psi_{t-4}}\cdots e^{\psi_1}) + (e^{\psi_t}e^{\psi_{t-1}}\cdots e^{\psi_1}). \qquad (24)$$

Recalling that for $n \geqslant 1$

$$e^{\psi_n} = \left(\frac{n+1}{n}\right)^{-\gamma}\frac{\tilde{h}_n}{\tilde{h}_{n-1}}$$

and for $n \geqslant 0$,

$$e^{\psi_0^n} = e^{a_0}\frac{\tilde{h}_0}{\tilde{h}_n}$$

we calculate for the coefficient in (24) of, for instance, A_{t-4}:

$$e^{\psi(01110)}e^{\psi(1110)}e^{\psi(110)}e^{\psi(10)} = e^{\psi_0^3}e^{\psi_3}e^{\psi_2}e^{\psi_1}$$
$$= \left(e^{a_0}\frac{\tilde{h}_0}{\tilde{h}_3}\right)\left(\frac{4}{3}\right)^{-\gamma}\frac{\tilde{h}_3}{\tilde{h}_2}\left(\frac{3}{2}\right)^{-\gamma}\frac{\tilde{h}_2}{\tilde{h}_1}\left(\frac{2}{1}\right)^{-\gamma}\frac{\tilde{h}_1}{\tilde{h}_0} = e^{a_0}4^{-\gamma}. \qquad (25)$$

To calculate the last term of (24), note that since $\tilde{h}_0 = e^{-a_0}$, and

$$\tilde{h}_n = (n+1)^{\gamma}\sum_{i=n+1}^{\infty} i^{-\gamma}$$

$$e^{\psi_n}e^{\psi_{n-1}}\cdots e^{\psi_1} = (n+1)^{-\gamma}\frac{\tilde{h}_n}{\tilde{h}_0} = e^{a_0}\sum_{i=n+1}^{\infty} i^{-\gamma}.$$

Therefore, continuing the equation (24) above we have

$$A_t = A_{t-1}e^{a_0}1^{-\gamma} + A_{t-2}e^{a_0}2^{-\gamma} + A_{t-3}e^{a_0}3^{-\gamma} + \cdots + A_1e^{a_0}(t-1)^{-\gamma} + e^{a_0}\sum_{i=t+1}^{\infty} i^{-\gamma}. \tag{26}$$

This is valid for all $t \geqslant 1$; we note that, in particular, for $t = 1$ one has only the final term, and so

$$A_1 = e^{\psi_1} = e^{a_0}\sum_{i=2}^{\infty} i^{-\gamma} = \mathcal{L}_{\psi}^1 I_{[1]}(z)$$

z in M_0.

Equation (26) is a recurrence relation; thus, since $A_0 = 0$ and A_1 are specified (as above), A_t is then determined recursively for all $t > 1$.

Formulae of this type come up in Tauberian theory, with applications to complex analysis, harmonic analysis, number theory and renewal theory. The latter is part of probability theory, which deals with counting the number of events when one has an independent identically distributed positive waiting time between occurrences of the events. What we want is to describe the asymptotic behaviour of A_t for $t \geqslant 1$. Standard results from this theory will help us in carrying out our calculation. For background see [KT75, GS92, Fel49, Fel71, Bar, Bre68, Pos79].

For $\gamma > 1$, now consider the probability measure on the line with distribution function F, defined by

$$dF = e^{a_0}\sum_{i=1}^{\infty} i^{-\gamma}\delta_i = \frac{1}{\zeta(\gamma)}\sum_{i=1}^{\infty} i^{-\gamma}\delta_i \tag{27}$$

where δ_i is point mass (the Dirac delta) at i (and where the measure is written above in Stieltjes form as dF). In the probabilistic interpretation, dF is the distribution of waiting times.

We set

$$f_i = \frac{i^{-\gamma}}{\zeta(\gamma)} \qquad i \geqslant 0. \tag{28}$$

The expected waiting time is the expected value of this probability measure which we write as

$$E \equiv \sum_{i=1}^{\infty} \mathrm{i} f_i = \frac{1}{\zeta(\gamma)}\sum_{i=1}^{\infty} i^{1-\gamma} = \frac{\zeta(\gamma-1)}{\zeta(\gamma)} \equiv u^{-1} > 1 \tag{29}$$

with the last equality coming from (9).

We write, for $t \geqslant 1$, the last term in (26) as a_t; thus

$$a_t \equiv \frac{1}{\zeta(\gamma)}\sum_{i=t+1}^{\infty} i^{-\gamma}. \tag{30}$$

From (6), we have for all t, and all $\gamma > 2$:

$$\frac{(t+2)^{1-\gamma}}{\zeta(\gamma)(\gamma-1)} \leqslant a_t \leqslant \frac{(t+1)^{1-\gamma}}{\zeta(\gamma)(\gamma-1)} \leqslant (t+1)^{1-\gamma} \leqslant t^{1-\gamma}. \tag{31}$$

We note that

$$A_1 = \mathrm{e}^{a_0} \sum_{i=2}^{\infty} i^{-\gamma} = a_1.$$

Equations (24) and (26) can therefore be rewritten in the form

$$A_t = \sum_{i=1}^{t-1} A_{t-i} \frac{1}{\zeta(\gamma)} i^{-\gamma} + a_t = a_t + \int_{-\infty}^{\infty} A_{t-x}\, \mathrm{d}F(x)$$

where we take $A_t = 0$ and $a_t = 0$ for $t \leqslant 0$. (The limits of integration could of course be replaced by 1 and $t - 1$.) This is an *arithmetic renewal equation*; it can be written as a convolution equation: $A = a + A * \mathrm{d}F$. Here a and $\mathrm{d}F$ are fixed and one wishes to solve for A, i.e. to find a fixed point for this affine map on sequence space. See also equation (40) below.

The solution of this equation in renewal theory proceeds as follows. Defining F_j to be the distribution function of the jth convolution of F with itself, i.e. the distribution of the independent random walk generated by F after j steps, one defines the *renewal function*

$$W_n \equiv \sum_{j=0}^{\infty} F_j(n).$$

Another interpretation of W_n is that it is the expected number of events up to time n [GS92, KT75]. One then shows that W is the unique solution to a second renewal equation:

$$W = F + W * \mathrm{d}F. \tag{32}$$

In our case, $W_0 = 0$ and W_n for $n > 0$ solves

$$W_n = \frac{1}{\zeta(\gamma)} \sum_{i=1}^{n} i^{-\gamma} + \frac{1}{\zeta(\gamma)} \sum_{i=1}^{n} W_{n-i} i^{-\gamma}.$$

(This is, again, a recurrence relation, so the specification of W_0 determines W_n for $n \geqslant 1$.)

For $\gamma > 2$, then the sum $\sum_{t=1}^{\infty} a_t$ is finite so theorem 5.1 in [KT75] applies. Therefore, the solution to the related equation above, $A = a + A * \mathrm{d}F$, is given by $A = a + a * \mathrm{d}W$, i.e. for integers t (recall $A_t = 0$ for $t \leqslant 0$):

$$A_t = a_t + \sum_{i=1}^{t-1} a_{t-i}\, \mathrm{d}W_i = a_t + \sum_{i=1}^{t-1} a_{t-i}(W_i - W_{i-1}). \tag{33}$$

Now one sees (cf [KT75]) that

$$\lim_{t\to\infty} A_t = \frac{1}{E} \sum_{j=1}^{\infty} a_j \tag{34}$$

which by definition equals

$$\frac{1}{E} \sum_{j=1}^{\infty} \left(\mathrm{e}^{a_0} \sum_{i=j+1}^{\infty} i^{-\gamma} \right) = \frac{\mathrm{e}^{a_0}}{E} \sum_{j=1}^{\infty} j(j+1)^{-\gamma}. \tag{35}$$

Meanwhile, from (10), we have

$$\mu[1] = \sum_{j=1}^{\infty} \mu([\underbrace{11\ldots1}_{j}0]) \equiv u \sum_{j=1}^{\infty}\sum_{i=j}^{\infty} \nu_i \equiv \frac{e^{a_0}}{E} \sum_{j=1}^{\infty} j(j+1)^{-\gamma}. \tag{36}$$

So, combining (34)–(36),

$$\mu[1] = \lim_{t\to\infty} A_t. \tag{37}$$

Remark 2.2. One part of the renewal theorem states that $\lim_{i\to\infty} W_{i+1} - W_i = \frac{1}{E} = u$. To estimate the difference $\mu[1] - A_t$ we shall need a precise estimate of this rate of convergence. Note first that

$$0 \leqslant (W_{i+1} - W_i) - \frac{1}{E} \tag{38}$$

(see pp 191–2 in [KT75]). Moreover, making use of a theorem of Nagaev, theorem 4′, p 98 in [Pos79], one can prove this stronger statement: $\exists \bar{c}_0 > 0$ such that for all i,

$$0 < (W_{i+1} - W_i) - \frac{1}{E} < \bar{c}_0 i^{(2-\gamma)^+}. \tag{39}$$

This last expression can then be used to give a proof of part (a) for the case $z \in M_0$. Indeed, this argument can be extended to a statement like (a) for a more general class of potentials similar in nature to ours. However, this approach using Nagaev's estimate (39) is not enough for the lower bound in part (b). Therefore, we will proceed in another way which will yield directly both the upper and lower estimates, for $z \in M_0$.

Thus, at this point we move on to the proof of (b).

Proof of theorem 2.6, part (b). We mention that our proof of the lower bound will, in a very strong way, use the special form of our potential, that is, of the sequence f_i defined in (28) above.

For this part we make the assumption that $\gamma > 2$. We summarize the definitions from above that we need, together with some basic properties of our example.

The sequence a_t defined in (30) satisfies

$$a_t = \frac{1}{\sum_{n=1}^{\infty} n^{-\gamma}} \sum_{j=t+1}^{\infty} j^{-\gamma} \qquad t = 1, 2, \ldots$$

is positive, summable and monotonically decreasing. By convention we take $a_0 = 0$.

The sequence f_n from (28) is

$$\frac{n^{-\gamma}}{\sum_{j=1}^{\infty} j^{-\gamma}} \qquad \gamma > 2 \quad n > 0$$

is also positive, summable and is monotonically decreasing; it is a probability sequence, with finite expected value: $\sum_{n=1}^{\infty} t f_t < \infty$. By convention $f_0 = 0$.

We observe that for the sequence $A_t = \mathcal{L}_{\psi}^t I_{[1]}(z),\ z \in M_0$ defined following (23), we have $A_0 = a_0 = 0$, $A_1 = a_1$; for $t > 1$, A_t satisfies the renewal equation

$$A_t = a_t + A_{t-1} f_1 + A_{t-2} f_2 + \cdots + A_0 f_t. \tag{40}$$

Let W_t be the solution of the renewal equation (32): then $W_0 = 0$, $W_1 = f_1$ and for $t > 1$

$$W_t = \left(\sum_{j=1}^{t} f_j\right) + W_{t-1} f_1 + W_{t-2} f_2 + \cdots + W_0 f_t.$$

From [KT75], p 184, theorem 4.1 and p 191, theorem 5.1 we have that $A_1 = a_1 + A_0 f_1 = a_1$ and for $t > 1$, the term A_t can be written as

$$A_t = a_t + a_{t-1}(W_1 - W_0) + a_{t-2}(W_2 - W_1) + \cdots + a_1(W_{t-1} - W_{t-2})$$

and A_t converges to

$$\frac{\sum_{n=1}^{\infty} a_t}{\sum_{n=1}^{\infty} t f_t} = \mu[1].$$

Next, we shall show that A_t is increasing; this will be essential for the lower bound estimation in part (b) of the theorem.

Lemma 2.7. *The A_t are positive and for $t > 1$ monotonically increasing in t.*

Proof. First, note that A_t is positive by induction in the renewal equation (40).

Now we will show monotonicity of A_t. Consider $B_t = A_{t+1} - A_t$ for $t \geqslant 0$. Remember that $f_0 = 0$, $a_0 = 0$, $A_0 = 0$, $W_0 = 0$ and $A_1 = a_1$ (therefore $B_0 = a_1$).

It follows from equation (40) (considering the difference $A_{t+1} - A_t$) that for $t > 0$ the B_t satisfy

$$B_t = (a_{t+1} - a_t) + B_{t-1} f_1 + B_{t-2} f_2 + \cdots + B_1 f_{t-1} + B_0 f_t.$$

Note that for $t > 0$, $a_{t+1} - a_t = -f_{t+1}$.

Define $b_t = (a_{t+1} - a_t) = -f_{t+1} < 0$, for $t > 0$ and $b_0 = a_1$. Then B_t satisfies for $t > 0$

$$B_t = b_t + B_{t-1} f_1 + B_{t-2} f_2 + \cdots + B_1 f_{t-1} + B_0 f_t \tag{41}$$

and $B_0 = b_0 = a_1$.

We want to show that $B_n > 0$, for all n.

Consider now the following power series in the complex variable z for $|z| < 1$:

$$\mathcal{B}(z) = \sum_{n=0}^{\infty} B_n z^n \qquad \text{and} \qquad f(z) = \sum_{n=1}^{\infty} f_n z^n.$$

From relation (41) and $b_t = -f_{t+1}$ it follows that

$$\mathcal{B}(z) = a_1 + \frac{f_1 z - f(z)}{z} + \mathcal{B}(z) f(z).$$

Therefore,

$$\mathcal{B}(z) = \frac{1}{1 - f(z)} \left(a_1 + \frac{f_1 z - f(z)}{z} \right).$$

We define c_n to be the coefficients of the power series

$$\frac{1}{1 - f(z)} = \sum_{n=0}^{\infty} c_n z^n. \tag{42}$$

Using the fact that

$$(1 - f(z)) \sum_{n=0}^{\infty} c_n z^n = \left(1 - \sum_{n=1}^{\infty} f_n z^n \right) \sum_{n=0}^{\infty} c_n z^n = 1$$

we obtain for $n > 0$ the relations $c_0 = 1$,

$$c_n - \sum_{j=1}^{n} f_j c_{n-j} = 0$$

or equivalently

$$-\sum_{j=2}^{n} f_j c_{n-j} = f_1 c_{n-1} - c_n.$$

Now we compute the sum

$$\begin{aligned}
\mathcal{B}(z) &= \sum_{n=0}^{\infty} c_n z^n \left(a_1 + \frac{f_1 z - f(z)}{z}\right) = \sum_{n=0}^{\infty} c_n z^n \left(f_2 - f_2 z + f_3 - f_3 z^2 + f_4 + f_4 z^3 + \cdots\right)\\
&= c_0(f_2 + f_3 + \cdots) + \big(c_1(f_2 + f_3 + \cdots) - c_0 f_2\big)\, z\\
&\quad + \big(c_2(f_2 + f_3 + \cdots) - (c_1 f_2 + c_0 f_3)\big)\, z^2 + \cdots\\
&\quad + \left(c_{n-1}(f_2 + f_3 + \cdots) - \left(\sum_{j=2}^{n} f_j c_{n-j}\right)\right) z^{n-1} + \cdots.
\end{aligned}$$

Using the fact that $\sum_{j=1}^{\infty} f_j = 1$, the last expression is

$$\begin{aligned}
\mathcal{B}(z) &= (f_2 + f_3 + \cdots) + \big(c_2(f_2 + f_3 + \cdots) + c_2 f_1 - c_3\big)\, z + \big(c_3(f_2 + f_3 + \cdots)\\
&\quad + (c_3 f_1 - c_4)\big)\, z^2 + \cdots + \big(c_n(f_2 + f_3 + \cdots) + (c_n f_1 - c_{n+1})\big)\, z^{n-1} + \cdots\\
&= (f_2 + f_3 + \cdots) + \big(c_2 - c_3\big)\, z + \big(c_3 - c_4)\big)\, z^2 + \cdots \big(c_{n-1} - c_n\big)\, z^{n-1} + \cdots.
\end{aligned}$$

The last expression gives a new form for the coefficients B_n of the power series for $\mathcal{B}(z)$. Thus in order to demonstrate that $B_n > 0$ for all n all we have to show is that the sequence c_n is monotonically decreasing.

It follows from a statement in Brietzke [Bri00] (using [Kal28]) that for the sequence f_j as above, the coefficients c_n of the power series in (42) are decreasing.

This result follows from a more general result: if $\sum_{j=1}^{\infty} f_j = 1$ and if the sequence f_j extends to some function F defined on the reals so that $f_j = F(j)$ is > 0 and decreasing and such that $\log F(x)$ is convex, and if also $f_2 \leqslant (1 - f_1) f_1$, then the above claim is true, that is the c_n are decreasing. Taking $F(x) = x^{-\gamma}$ gives this result for our example (see [Bri00] for a full explanation).

Finally, from the fact that the c_n are monotonically decreasing we conclude that the sequence A_t is monotonically increasing, concluding the proof of the lemma. □

Now we recall what we wish to prove, part (b) in theorem 2.6: that there exist $\overline{C}_1, C_1 > 0$ such that $\overline{C}_1 n^{2-\gamma} \leqslant \mu[1] - A_n \leqslant C_1 n^{2-\gamma}$. To prove this we shall make use of two classical theorems of Fourier analysis; see p 230 of [Bar], vol II and also [Har49].

Theorem A. *Suppose $g(\theta) = \sum_{n=1}^{\infty} b_n \sin(n\,\theta)$, where b_n converges to zero. If $a > 0$ and $g(\theta)$ is Hölder of order a then there exists a constant c such that $b_n < cn^{-(1+a)}$.*

This has a converse:

Theorem B. *Suppose $g(\theta) = \sum_{n=1}^{\infty} b_n \sin(n\,\theta)$, where b_n is monotonically decreasing to zero. If $a > 0$ and there exists a constant c such that $b_n < cn^{-(1+a)}$, then $g(\theta)$ is Hölder of order a.*

We will also need a theorem from [Pos79, p 33].

Theorem C. *Let $\sum_{n=1}^{\infty} b_n z^n$ be a power series converging for $|z| < 1$. Suppose that the coefficients b_n are non-negative and non-increasing. Suppose in addition that*

$$\sum_{n=1}^{\infty} b_n z^n = \frac{1}{(1-z)^{\beta}} + \mathrm{o}\left(\frac{1}{(1-t)^{\beta}}\right)$$

for some β, $0 < \beta < 1$. Then

$$\lim_{n\to\infty} \Gamma(\beta) n^{1-\beta} b_n = 1.$$

There is a converse of theorem C as well (see theorem 2, p 12 [Pos79]).

Theorem D. *Let $\delta > 0$, A be constants and let*

$$S_N = \sum_{n=1}^{N} b_n \approx \frac{A\,N^{\delta}}{\Gamma(1+\delta)}$$

as N goes to ∞, then

$$\lim_{x\to 1-0} (1-z)^{\delta}\left(\sum_{n=1}^{\infty} b_n z^n\right) = A.$$

We note that by taking $z = \mathrm{e}^{\mathrm{i}\theta}$ in the last two theorems one sees the relationship to the first two; the value $\theta = 0$ corresponds to $z = 1$. Theorems A–D are Tauberian theorems, which describe different aspects of the behaviour of the series $\sum_{n=1}^{\infty} b_n z^n$ for z close to 1, in terms of the rate of convergence of the b_n to zero.

Setting $V_t = \mu[1] - A_t$, then from the renewal equation (40), it follows that

$$V_t = \mu[1](f_t + f_{t+1} + \cdots) - (f_{t+1} + f_{t+2} + \cdots) + \sum_{j=1}^{t-1} V_{t-j} f_j.$$

Note that since A_t is monotonic, so is V_t.

Setting $K_t = \mu[1](f_t + f_{t+1} + \cdots) - (f_{t+1} + f_{t+2} + \cdots)$, the last equation can be written as

$$V_t = K_t + \sum_{j=1}^{t-1} V_{t-j} f_j.$$

Note that K_t goes to zero with order $t^{1-\gamma}$, since $\mu[1] < 1$. Moreover, $S_N = \sum_{t=1}^{N} K_n \approx C N^{\gamma-2}$.

Consider the power series

$$V(z) = \sum_{n=1}^{\infty} V_n z^n \qquad f(z) = \sum_{n=1}^{\infty} f_n z^n \qquad \text{and} \qquad K(z) = \sum_{n=1}^{\infty} K_n z^n.$$

From the equation of the V_t above

$$V(z) = K(z) + f(z)V(z)$$

and therefore,

$$V(z) = \frac{K(z)}{1 - f(z)}.$$

We study first the case $3 > \gamma > 2$.

Considering now $z = e^{i\theta} = \cos\theta + i\sin\theta$, $b_n = nf_n$ and $a = \gamma - 2$ in theorems A and B, then

$$f'(z) = \sum_{n=1}^{\infty} nf_n z^{n-1}$$

exists and is Hölder of class a. We know from theorem B (taking there $b_n = K_n$) that $K(z)$ is Hölder of class $\gamma - 2$ since K_n is of the order of $n^{1-\gamma}$.

Dividing by $(1-z)^{\gamma-3}$ in both sides of the last equation we obtain

$$\frac{V(z)}{(1-z)^{\gamma-3}} = \frac{K(z)}{(1-z)^{\gamma-2}} \frac{1-z}{1-f(z)}.$$

It is important to note that these limits exist:

$$\lim_{z\to 1^-} \frac{1-z}{1-f(z)} \qquad \text{and} \qquad \lim_{z\to 1^-} \frac{K(z)}{(1-z)^{\gamma-2}}.$$

The first is due to the fact that f is differentiable at 1 (see [Bar]) and the second follows from theorem D. We note that $S_N = \sum_{t=1}^{N} K_n \approx CN^{\gamma-2}$.

From theorem C we see that for $2 < \gamma < 3$, taking $\beta = 3-\gamma$ and $b_n = V_n$, the coefficients $V_n = \mu[1] - A_n$ satisfy the condition we wish to prove, of part (b) in theorem 2.6. Indeed, the limit of

$$\frac{\mu[1] - A_n}{n^{2-\gamma}} = \frac{V_n}{n^{2-\gamma}}$$

as $n \to \infty$ exists (this follows from considering the series $\sum_{n=1}^{\infty} dV_n z^n$ with a suitable constant d and applying theorem C). We emphasize the important role played here by the monotonicity of the V_n (which followed from that of the A_n); this guaranteed the hypothesis of theorem C.

For the case $\gamma > 3$ the last relation is also true. This is obtained using the fact that $K(z)$ is differentiable of class $C^{\gamma-2}$. Taking the k-derivative of $K(z)$, where k is the integer part of $\gamma - 2$, we obtain a $(\gamma - 2 - k)$-Hölder function and we apply a similar argument to that used before.

This concludes the proof of part (b) of theorem 2.6. □

We shall continue the proof of theorem 2.6 after we show how part (b) of the theorem can be applied to prove lower as well as upper decay of autocorrelation estimates for a specific function, and show the relation of this to '$1/f$-noise'; see also the explanation and definition given in the introduction.

Theorem 2.8. (Lower and upper bounds for decay of autocorrelation and '$1/f$-noise' phenomenon for $I_{[0]}$).
(a) For $\gamma > 2$ and $\varphi(x) = \phi(x) = I_{[0]}(x)$, there exist $c, C > 0$ such that

$$c\,t^{2-\gamma} \leqslant \int (\varphi \circ T^t)\,\phi\,d\mu - \int \varphi\,d\mu \int \phi\,d\mu \leqslant C\,t^{2-\gamma}. \tag{43}$$

(b) For $\gamma \in (2, 3)$, the sequence of random variables $\phi \circ T^t$ exhibits $1/f^{3-\gamma}$-noise.

Proof. Recall that $A_t = \mathcal{L}^t I_{[1]}(z)$ for $z \in M_0$, where this function is constant. We define $X_t = \mathcal{L}_\psi^t I_{[0]}(z)$ for $z \in M_0$ (where, similarly, this is constant). Next note that $\mathcal{L}^t I_{[0]} + \mathcal{L}^t I_{[1]} = \mathcal{L}^t I_{\Sigma^+} \equiv 1$ and $\mu[0] + \mu[1] = 1$. Therefore, the rate of convergence of A_t to $\mu([1])$ is the same as the rate of convergence of X_t to $\mu([0])$. That is, the monotonicity of the sequence of numbers A_t implies that of the X_t, and from part (b) of theorem 2.6 we have:

$$\overline{C}_1 t^{2-\gamma} \leqslant X_t - \mu[0] \leqslant C_1 t^{2-\gamma}. \tag{44}$$

By proposition 6.1 below we have that $\int (\varphi \circ T^t)\phi \, \mathrm{d}\mu = \int \varphi \, \mathcal{L}_\psi^t(\phi) \, \mathrm{d}\mu$.

Therefore, we have

$$\int (\varphi \circ T^t)\phi \, \mathrm{d}\mu = \int \phi \, \mathcal{L}_\psi^t(\phi) \, \mathrm{d}\mu = \int_{[0]} \mathcal{L}_\psi^t(\phi) \, \mathrm{d}\mu = \int_{[0]} X_t \, \mathrm{d}\mu = \mu([0]) \cdot X_t \tag{45}$$

and so:

$$\int (\varphi \circ T^t)\phi \, \mathrm{d}\mu - \int \varphi \, \mathrm{d}\mu \int \phi \, \mathrm{d}\mu = \mu([0]) \cdot \big(X_t - \mu([0])\big). \tag{46}$$

Hence from (44), we have proved part (a), taking c, C to be $\mu([0])$ times $\overline{C}_1, C_1$.

It follows from this that the spectral measure (see theorem 7.3 and definitions in section 9.7 in [KT75]),

$$F(\lambda) = \sum_{t=0}^{\infty} \left(\int (\phi \circ T^t)\phi \, \mathrm{d}\mu - \int \phi \, \mathrm{d}\mu \int \phi \, \mathrm{d}\mu \right) \cos(t\,\lambda)$$

then (by the theorems A–D above) F is Hölder for $\gamma > 3$ and $F(\lambda) \approx \lambda^{\gamma-3}$, when $\lambda \approx 0$, for $3 > \gamma > 2$ (that is, there exists c, C such that $c\lambda^{\gamma-3} \leqslant F(\lambda) \leqslant C\lambda^{\gamma-3}$). For this last statement apply theorem D above, with $z = \mathrm{e}^{\mathrm{i}t\lambda}$. Note that $z = 1$ corresponds to $\lambda = 0$. □

We mention that a similar result proved using other methods appears on p 163 in [Lop93].

Now we return to the proof of theorem 2.6, part (a). For the rest of the proof (and the rest of the paper) we shall have no further need for the lower bound given in part (b).

Proof of theorem 2.6 (The case $z \in M_s$ for $s \geqslant 1$). Next we will consider z of the form $z = (\underbrace{11\ldots1}_{s}0z_{s+1}\ldots)$, i.e. where $z \in M_s$, for $s \geqslant 1$. We know from [Hof77] that for $\gamma > 2$, writing

$$A_t^s \equiv \mathcal{L}_\psi^t I_{[1]}(z) \tag{47}$$

then A_t^s converges to $\mu[1]$ as $t \to \infty$; we will show more precisely that there exists a constant $\hat{c}_\gamma$ such that for z as above,

$$|\mu[1] - A_t^s| \leqslant \hat{c}_\gamma \left((s+t)^{2-\gamma} + \left(\frac{s+t}{s}\right)^{1-\gamma} \right).$$

Recall that we are writing $\psi_0^n \equiv \psi(0\underbrace{11\ldots1}_{n}0)$, $\psi_n \equiv \psi(\underbrace{11\ldots1}_{n}0)$. From figure 3 it follows that

$$A_t^s = A_{t-1}\mathrm{e}^{\psi_0^s} + A_{t-2}(\mathrm{e}^{\psi_0^{s+1}}\mathrm{e}^{\psi_{s+1}}) + A_{t-3}(\mathrm{e}^{\psi_0^{s+2}}\mathrm{e}^{\psi_{s+2}}\mathrm{e}^{\psi_{s+1}}) + \cdots$$
$$+ A_1(\mathrm{e}^{\psi_0^{t+s-2}}\mathrm{e}^{\psi_{t+s-2}}\ldots\mathrm{e}^{\psi_{s+1}}) + (\mathrm{e}^{\psi_{s+t}}\mathrm{e}^{\psi_{s+t-1}}\ldots\mathrm{e}^{\psi_{s+1}}).$$

The coefficient for A_{t-1} is, from (17),

$$\mathrm{e}^{\psi_0^s} = \mathrm{e}^{a_0}\frac{\tilde{h}_0}{\tilde{h}_s} = \frac{1}{\tilde{h}_s} = \left(\frac{s+1}{s+1}\right)^{-\gamma}\frac{1}{\tilde{h}_s}.$$

So by (16) and (17) we can write the above expression as

$$
\begin{aligned}
A_t^s &= A_{t-1}\left(\frac{s+1}{s+1}\right)^{-\gamma}\frac{1}{\tilde{h}_s} + A_{t-2}\left(\frac{s+2}{s+1}\right)^{-\gamma}\frac{1}{\tilde{h}_s} + \cdots + A_1\left(\frac{s+t-1}{s+1}\right)^{-\gamma}\frac{1}{\tilde{h}_s} \\
&\quad + \left(\frac{s+t+1}{s+1}\right)^{-\gamma}\frac{\tilde{h}_{s+t}}{\tilde{h}_s} \\
&= \frac{(s+1)^{\gamma}}{\tilde{h}_s}\left(A_{t-1}(s+1)^{-\gamma} + \cdots + A_1(s+t-1)^{-\gamma} + \frac{\tilde{h}_{s+t}}{(s+t+1)^{\gamma}}\right). \qquad (48)
\end{aligned}
$$

Recalling from (8) that

$$\tilde{h}_n = (n+1)^{\gamma}\sum_{i=n+1}^{\infty} i^{-\gamma}$$

we have

$$
\begin{aligned}
A_t^s &= \frac{1}{\sum_{i=s+1}^{\infty} i^{-\gamma}}\left(A_{t-1}(s+1)^{-\gamma} + \cdots + A_1(s+t-1)^{-\gamma} + \sum_{i=s+t+1}^{\infty} i^{-\gamma}\right) \\
&= \frac{1}{\sum_{i=s+1}^{\infty} i^{-\gamma}}\left(\sum_{j=1}^{t-1} A_{t-j}(s+j)^{-\gamma} + \sum_{i=s+t+1}^{\infty} i^{-\gamma}\right). \qquad (49)
\end{aligned}
$$

Now

$$
\begin{aligned}
\mu[1] - (\mathcal{L}_{\psi}^{t} I_{[1]})(z) &= \lim_{j\to\infty} A_j^s - A_t^s = \mu[1] - A_t^s \\
&= \left(\mu[1] - \frac{1}{\sum_{i=s+1}^{\infty} i^{-\gamma}}\sum_{j=1}^{t-1} A_{t-j}(s+j)^{-\gamma}\right) - \frac{\sum_{i=s+t+1}^{\infty} i^{-\gamma}}{\sum_{i=s+1}^{\infty} i^{-\gamma}}. \qquad (50)
\end{aligned}
$$

To study the first part of (50), recall from part (b) of theorem 2.6 that $\overline{C}_1 \cdot t^{2-\gamma} \leqslant \mu[1] - A_t \leqslant C_1 t^{2-\gamma}$.

We define c_t by the equation $\mu[1] - A_t = c_t t^{2-\gamma}$; note that therefore $0 < \overline{C}_1 < c_t \leqslant C_1$ for all such t; in what follows we shall only make use of the upper bound here. We will substitute $\mu[1] - c_k k^{2-\gamma}$ for A_k in the first part of the expression (50). Thus we have for this first part:

$$
\begin{aligned}
\mu[1] - \frac{1}{\sum_{i=s+1}^{\infty} i^{-\gamma}}\left(\sum_{j=1}^{t-1} A_{t-j}(s+j)^{-\gamma}\right) &= \left(\mu[1] - \frac{1}{\sum_{i=s+1}^{\infty} i^{-\gamma}}\sum_{j=1}^{t-1}\mu[1](s+j)^{-\gamma}\right) \\
&\quad + \frac{1}{\sum_{i=s+1}^{\infty} i^{-\gamma}}\sum_{j=1}^{t-1} c_{t-j}(t-j)^{2-\gamma}(s+j)^{-\gamma}.
\end{aligned}
$$

The first term here is

$$= \frac{\mu[1]}{\sum_{i=s+1}^{\infty} i^{-\gamma}}\left(\sum_{i=s+1}^{\infty} i^{-\gamma} - \sum_{i=s+1}^{s+t-1} i^{-\gamma}\right) = \mu[1]\cdot\frac{\sum_{i=s+t}^{\infty} i^{-\gamma}}{\sum_{i=s+1}^{\infty} i^{-\gamma}}. \qquad (51)$$

Thus the quantity we want to estimate from above, equation (50), is equal to the modulus of

$$\mu[1]\cdot\frac{\sum_{i=s+t}^{\infty} i^{-\gamma}}{\sum_{i=s+1}^{\infty} i^{-\gamma}} - \frac{\sum_{i=s+t+1}^{\infty} i^{-\gamma}}{\sum_{i=s+1}^{\infty} i^{-\gamma}} + \frac{1}{\sum_{i=s+1}^{\infty} i^{-\gamma}}\sum_{j=1}^{t-1} c_{t-j}(t-j)^{2-\gamma}(s+j)^{-\gamma}. \qquad (52)$$

Let us write this as (a) − (b) + (c). First we analyse terms (a) and (b).

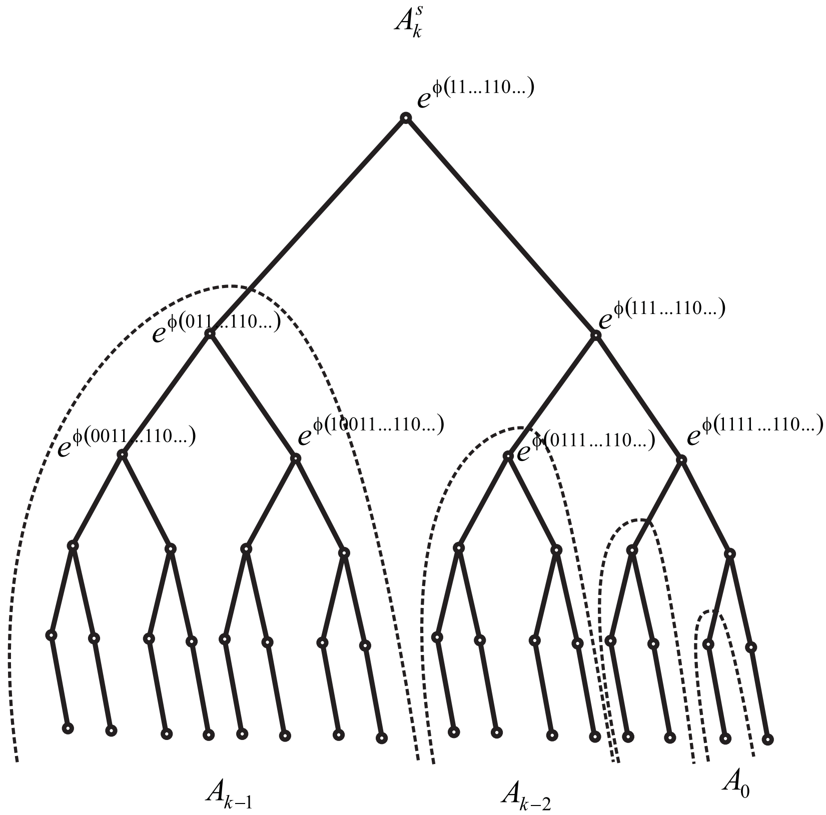

Figure 3. Geometric visualization for the equation of A_k^s.

Note that since s and $t \geqslant 1$,

$$\frac{s+t+2}{s+1} \leqslant 2\frac{s+t}{s} \qquad \text{and} \qquad \frac{s}{s+2} \geqslant \frac{1}{3} \tag{53}$$

so for the lower and upper bounds of (a) and (b) we have, using (53):

$$2^{1-\gamma}\left(\frac{s+t}{s}\right)^{1-\gamma} \leqslant \left(\frac{s+t+1}{s+1}\right)^{1-\gamma} \leqslant \frac{(\mathrm{a})}{\mu[1]} \leqslant \left(\frac{s+t}{s+2}\right)^{1-\gamma} \leqslant 3^{\gamma-1}\left(\frac{s+t}{s}\right)^{1-\gamma} \tag{54}$$

and

$$2^{1-\gamma}\left(\frac{s+t}{s}\right)^{1-\gamma} \leqslant \left(\frac{s+t+2}{s+1}\right)^{1-\gamma} \leqslant (\mathrm{b}) \leqslant \left(\frac{s+t+1}{s+2}\right)^{1-\gamma} \leqslant 3^{\gamma-1}\left(\frac{s+t}{s}\right)^{1-\gamma}. \tag{55}$$

Thus for the difference (a) − (b) we have

$$2^{1-\gamma}(\mu[1]-1)\left(\frac{s+t}{s}\right)^{1-\gamma} \leqslant (\mathrm{a}) - (\mathrm{b}) \leqslant 3^{\gamma-1}(\mu[1]-1)\left(\frac{s+t}{s}\right)^{1-\gamma}. \tag{56}$$

The last term of (52) is

$$(\mathrm{c}) \equiv \frac{1}{\sum_{i=s+1}^{\infty} i^{-\gamma}} \sum_{j=1}^{t-1} c_{t-j}(t-j)^{2-\gamma}(s+j)^{-\gamma} \tag{57}$$

which has, by (6), an upper bound of

$$\text{(c)} \leqslant C_1 \frac{(\gamma - 1)}{(s+2)^{1-\gamma}} \sum_{j=1}^{t-1} (t-j)^{2-\gamma}(s+j)^{-\gamma}. \tag{58}$$

Applying lemma A4 from the appendix with $\bar{c}_3 \equiv \hat{c}_{\alpha,\beta}$ for $\alpha = -\gamma$, $\beta = 2 - \gamma$ we have, taking first the case $\gamma > 3$ (so $\beta < -1$), and setting $\bar{c}_5 = (\gamma - 1)\bar{c}_3 C_1$, that (c) is bounded above by

$$\frac{\bar{c}_5}{(s+2)^{1-\gamma}} \left(s^{1-\gamma}(s+t)^{2-\gamma} + (s+t)^{-\gamma} \right) \leqslant \bar{c}_5 \left(\left(\frac{s}{s+2} \right)^{1-\gamma} (s+t)^{2-\gamma} + \left(\frac{s+t}{s+2} \right)^{1-\gamma} \right). \tag{59}$$

Hence

$$\text{(c)} \leqslant c_5 \left((s+t)^{2-\gamma} + \left(\frac{s+t}{s} \right)^{1-\gamma} \right) \tag{60}$$

where $c_5 = \bar{c}_5 \cdot 3^{\gamma-1}$, using (53).

Next consider the case $\gamma = 3$, so $\beta = -1$. Then from lemma A4, using (60),

$$\begin{aligned} \text{(c)} &\leqslant \frac{\bar{c}_6}{s^{1-\gamma}} \left(s^{1-\gamma}(s+t)^{2-\gamma} + (s+t)^{-\gamma} \log(s+t) \right) \\ &= \bar{c}_6 \left((s+t)^{2-\gamma} + \frac{(s+t)^{-\gamma}}{s^{1-\gamma}} \log(s+t) \right) \leqslant \bar{c}_6 \left((s+t)^{2-\gamma} + \left(\frac{s+t}{s} \right)^{1-\gamma} \right) \end{aligned}$$

where $\bar{c}_6 = (\gamma - 1) \cdot 3^{\gamma-1} C_1$.

For $\gamma \in (2, 3)$, so $\beta \in (-1, 0)$, we have from lemma A4 that (c) satisfies

$$\text{(c)} \leqslant \frac{\bar{c}_6}{s^{1-\gamma}} \left(s^{1-\gamma}(s+t)^{2-\gamma} + (s+t)^{3-2\gamma} \right) \leqslant \bar{c}_6 \left((s+t)^{2-\gamma} + \left(\frac{s+t}{s} \right)^{1-\gamma} \right).$$

Combining this with (56) we therefore have these upper and lower bounds for (52):

$$\begin{aligned} 2^{1-\gamma}(\mu[1] - 1) \left(\frac{s+t}{s} \right)^{1-\gamma} &\leqslant \text{(a)} - \text{(b)} \leqslant \text{(a)} - \text{(b)} + \text{(c)} \\ &\leqslant \text{(a)} + \text{(c)} \leqslant \bar{c}_\gamma \left((s+t)^{2-\gamma} + \left(\frac{s+t}{s} \right)^{1-\gamma} \right) \end{aligned}$$

where $\bar{c}_\gamma = \max\{c_5, \bar{c}_6\} + 3^{\gamma-1}\mu[1]$.

Finally, therefore, for the modulus of (52) we have this upper bound:

$$|\mu[1] - (\mathcal{L}_\psi^t I_{[1]})(z)| \leqslant \hat{c}_\gamma \left((s+t)^{2-\gamma} + \left(\frac{s+t}{s} \right)^{1-\gamma} \right) \tag{61}$$

where $\hat{c}_\gamma = \max\{\bar{c}_\gamma, 2^{1-\gamma}(1 - \mu[1])\}$. This is valid for each $\gamma > 2$, for all $z \in M_s$, for all $s, t \geqslant 1$. This completes the proof of part (a) of theorem 2.6. □

Remark 2.3. For t fixed, when s goes to infinity $(s+t)^{2-\gamma}$ goes to zero and

$$\left(\frac{s+t}{s} \right)^{1-\gamma}$$

goes to 1. This indicates that uniform convergence should *not* hold true. (For a rigorous proof of this we would need to extend our lower bound estimates to other points z.)

Next we use the result of theorem 2.6 (for the zero-cylinder set $P = [1]$) to study arbitrary k-cylinder sets, for $k \geqslant 0$. Each $P = [w_0 \dots w_k] \in \mathcal{C}_k$ can be written uniquely in the form

$$P = [w_0 \dots w_{j-1} 0 \underbrace{11 \dots 1}_{m}] \tag{62}$$

where $j + m = k$, and where we allow $m = k + 1$, $j = 0$ (so $P = [\underbrace{11 \dots 1}_{k+1}]$), and also $m = 0$ (so $P = [w_0 \dots w_{k-1} 0]$). For general cylinders we have the following; see the preprint version at http://ime.usp.br/∼afisher for the proof.

Theorem 2.9. *Let $\gamma > 2$. There exist $\bar{c}, \bar{d} > 0$ such that for all k, for all $P \in \mathcal{C}_k$, with*

$$\begin{aligned} &P = [w_0 \dots w_{j-1} 0 \underbrace{11 \dots 1}_{m}] \\ &|\mathcal{L}_\psi^t I_P(z) - \mu P| \leqslant \bar{c} \frac{\mu[w_0 \dots w_{j-1} 0]}{\mu[0]} \big((s + t - k)^{(2-\gamma)^+} + (k + 1) B(s, t - k)\big) \end{aligned} \tag{63}$$

for all $t > k$ and for all $z \in M_s$, $s = 0, 1, 2, \dots$, where

$$B(0, l) = 0 \qquad \text{and} \qquad B(s, l) = \left(\frac{s + l}{s}\right)^{1-\gamma} \qquad \text{for} \quad s \geqslant 1.$$

3. Mixing and weak Bernouilli

Next we use theorem 2.9 to analyse the rate of mixing:

Theorem 3.1. *Let $\gamma > 2$, and let μ be the invariant measure defined from the potential g_γ. Then we have, for every $k \geqslant 1$,*

$$\sum_{P, Q \in \mathcal{C}_k} \left|\mu(P \cap T^{-t} Q) - \mu P \mu Q\right| \leqslant \bar{c}(k + 2)^2 (t - k)^{(2-\gamma)^+}.$$

Proof. The sum is taken over all k-cylinder sets, i.e. all $P, Q \in \mathcal{C}_k$. For summing over the sets P, we split $\mathcal{C}_k$ into $(k + 2)$ subcollections: all sets P of the form $[w_0 \dots w_{j-1} 0 \underbrace{11 \dots 1}_{m}]$, with $j + m = k + 1$, $m = 0, 1, 2, \dots, k + 1$.

For each m we denote this subcollection by $\mathcal{C}_{k,m}$; thus $\mathcal{C}_k = \cup_{m=0}^{k+1} \mathcal{C}_{k,m}$.

We note first that for j fixed, $\sum_{w_0 \dots w_{j-1}} \mu[w_0 \dots w_{j-1} 0] = \mu[\underbrace{* * * \dots * 0}_{j}] = \mu[0]$ since μ is invariant (where $*$ denotes 'no restrictions').

Now since the Ruelle operator is the dual of the Koopman operator (see proposition 6.1 of section 6), i.e. $U^* = \mathcal{L}$, for f, g measurable we have $\int \mathcal{L} f \cdot g \, \mathrm{d}\mu = \int f \cdot g \circ T \, \mathrm{d}\mu$. Therefore,

$$\begin{aligned} \mu(P \cap T^{-t} Q) - \mu P \mu Q &= \int I_P \cdot (I_Q \circ T^t) \, \mathrm{d}\mu - \mu P \int I_Q \, \mathrm{d}\mu \\ &= \int (\mathcal{L}_\psi^t I_P) \cdot I_Q - \mu P \cdot I_Q \, \mathrm{d}\mu = \int I_Q (\mathcal{L}_\psi^t I_P - \mu P) \, \mathrm{d}\mu. \end{aligned} \tag{64}$$

Now for $P \in \mathcal{C}_{k,m}$ and $z \in M_s$, $s \geqslant 0$, we know from (63) that

$$\left|\mathcal{L}_\psi^t I_P(z) - \mu P\right| \leqslant \bar{c} \, \frac{\mu[w_0 \dots w_{j-1} 0]}{\mu[0]} \big((s + t - k)^{(2-\gamma)^+} + (k + 1) B(s, t - k)\big) \tag{65}$$

$$\leqslant \bar{c} \, \frac{\mu[w_0 \dots w_{j-1} 0]}{\mu[0]} \big((t - k)^{(2-\gamma)^+} + (k + 1) B(s, t - k)\big). \tag{66}$$

Using this together with the estimate (11) (that is, for a constant z_γ, we have $\mu(M_s) \leqslant z_\gamma(s+1)^{1-\gamma}$) we conclude that for any fixed k-cylinder set P,

$$\int \left|\mathcal{L}_\psi^t I_P(z) - \mu P\right| \mathrm{d}\mu(z) \leqslant \sum_{s=0}^{\infty} \mu(M_s)\bar{c}\, \frac{\mu[w_0 \dots w_{j-1}0]}{\mu[0]} \times \big((t-k)^{(2-\gamma)^+} + (k+1)B(s, t-k)\big)$$

$$\leqslant \bar{c}\,(k+2)\frac{\mu[w_0 \dots w_{j-1}0]}{\mu[0]}\big(t-k\big)^{(2-\gamma)^+} \tag{67}$$

where we have used (6) to estimate the sum.

Putting these facts together, we have

$$\sum_{P,Q\in\mathcal{C}_k} \left|\mu(P\cap T^{-t}Q) - \mu P \mu Q\right| \leqslant \sum_{P,Q\in\mathcal{C}_k} \int_{\Sigma^+} I_Q \left|\mathcal{L}_\psi^t I_P - \mu P\right| \mathrm{d}\mu$$

$$= \sum_{P\in\mathcal{C}_k} \int_{\Sigma^+} \left|\mathcal{L}_\psi^t I_P - \mu P\right| \mathrm{d}\mu \tag{68}$$

$$\leqslant \sum_{m=0}^{k+1} \sum_{P\in\mathcal{C}_{k,m}} \bar{c}\,(k+2)\frac{\mu[w_0 \dots w_{j-1}0]}{\mu[0]}\big(t-k\big)^{(2-\gamma)^+}$$

$$= \bar{c}\,(k+2)^2\big(t-k\big)^{(2-\gamma)^+} \tag{69}$$

as claimed. □

We therefore have immediately the following analogues of the classical theorems from Bowen's book [Bow75]:

Corollary 3.2. *Let $\gamma > 2$. The shift map T with the measure μ is mixing*

and

Corollary 3.3. *T with μ and with the standard generating partition $\{[0],[1]\}$ is weak Bernouilli. Hence (by Ornstein's theorem) the transformation is measure-theoretically isomorphic to a Bernouilli shift of the same entropy.*

4. Polynomial decay of correlation for $\mathcal{SP}$ observables

We will use the following notation: $f(t) \leqslant c \cdot t^{a^{++}}$ means that for each $\delta > 0$, there exists $R_\delta > 0$ such that $f(t) \leqslant c \cdot t^{a+\delta}$ for all $t > R_\delta$. This is slightly weaker than $f(t) \leqslant c \cdot t^{a^+}$, where there are no lower restrictions on the variable t.

Theorem 4.1. *For $\gamma > 2$, there exists $C > 0$ such that for all $f, g \in \mathcal{SP}$,*

$$\left|\int f \cdot g\circ T^t \,\mathrm{d}\mu - \int f\,\mathrm{d}\mu \int g\,\mathrm{d}\mu\right| \leqslant C(\|f\|_{\mathcal{SP}} + \|g\|_{\mathcal{SP}})^2\, t^{(2-\gamma)^{++}}. \tag{70}$$

Proof. We adapt the proof of lemma 1.14 in [Bow75] (for Hölder functions and exponential decay) to the present situation. Let f_k denote the conditional expectation of f on the σ-algebra generated by the k-cylinder sets $\mathcal{C}_k$. That is, for $z \in P \in \mathcal{C}_k$,

$$f_k(z) \equiv \frac{1}{\mu P}\int_P f\,\mathrm{d}\mu.$$

Hence

$$\int_P f_k \, \mathrm{d}\mu = \int_P f \, \mathrm{d}\mu$$

for all $P \in \mathcal{C}_k$; and f_k is a step function, a k-cylinder approximation to f. Writing a_P for the value of f_k on P and b_Q for the value of g_k on Q, we have

$$f_k = \sum_{P \in \mathcal{C}_k} a_P I_P \qquad g_k = \sum_{Q \in \mathcal{C}_k} b_Q I_Q .$$

Since f is in $\mathcal{SP}$, by (1) we have $\|f - f_k\|_\infty \leqslant \mathrm{var}_k\, f \leqslant 2\|f\|_{\mathcal{SP}} k^{-a}$, for all $a > 0$ and for all $k > \underline{k_a}(f)$.

Now note first that (writing integrals $\int f \, \mathrm{d}\mu = \mu f$) we have, by theorem 3.1,

$$\begin{aligned} \left|\mu(f_k \cdot g_k \circ T^t) - \mu f_k \cdot \mu g_k\right| &= \Big| \sum_{P, Q \in \mathcal{C}_k} a_P b_Q (\mu(P \cap T^{-t} Q) - \mu P \mu Q) \Big| \\ &\leqslant \|f\|_\infty \|g\|_\infty \bar{c}\, (k+2)^2 \big(t-k\big)^{(2-\gamma)^+}. \end{aligned} \tag{71}$$

Hence we have

$$\begin{aligned} \left|\mu(f \cdot g \circ T^t) - \mu f \mu g\right| = \big|&(\mu(f_k \cdot g_k \circ T^t) - \mu f_k \cdot \mu g_k) + \mu((f - f_k) g \circ T^t)) \\ &+ \mu(f_k \cdot ((g - g_k) \circ T^t))\big|. \end{aligned}$$

For the last equality we have used the fact that $\mu(f - f_k) = \mu(g - g_k) = 0$. We write $\underline{C_a} = \max\{\underline{c_a}(f), \underline{c_a}(g)\}$ and $\underline{K_a} = \max\{\underline{k_a}(f), \underline{k_a}(g)\}$. Hence, using the definition of the $\mathcal{SP}$ norm and (71), this expression is

$$\begin{aligned} &\leqslant \left|\mu(f_k \cdot g_k \circ T^t) - \mu f_k \cdot \mu g_k\right| + \|f - f_k\|_\infty \cdot \|g \circ T^t\|_1 + \|g - g_k\|_\infty \cdot \|f_k\|_1 \\ &\leqslant \|f\|_\infty \|g\|_\infty \bar{c}\, (k+2)^2 \big(t-k\big)^{(2-\gamma)^+} + 2\underline{C_a} k^{-a} \|g\|_\infty + 2\underline{C_a} k^{-a} \|f\|_\infty \\ &\leqslant 2(\|f\|_{\mathcal{SP}} + \|g\|_{\mathcal{SP}})^2 \big(\bar{c}\, (k+2)^2 \big(t-k\big)^{(2-\gamma)^+} + k^{-a}\big). \end{aligned}$$

This holds for all $a > 0$ and for all t and k with $t > k > \underline{K_a}$.

Now choose $\delta \in (0, 1)$, and let $a > (2/\delta)((2-\gamma)^+)$. Then for $k = [t^{\delta/2}] + 1$ where $[\cdot]$ indicates the integer part, and $k > \underline{K_a}$, we have $k > t^{\delta/2}$, so $k^{-a} < t^{(2-\gamma)^+}$ and

$$\frac{k}{t} < \frac{t^{\delta/2} + 1}{t} = t^{\delta/2 - 1} + 1/t$$

which is $< \frac{1}{2}$ for $\delta < 1$ and $t > 9$.

Hence $(1 - k/t) > \frac{1}{2}$, so $(1 - k/t)^{(2-\gamma)^+} < \left(\frac{1}{2}\right)^{(2-\gamma)^+} = 2^{\gamma^+ - 2}$.

Therefore, $(t-k)^{(2-\gamma)^+} = \big((t-k)/t\big)^{(2-\gamma)^+} t^{(2-\gamma)^+} \leqslant 2^{\gamma^+ - 2} t^{(2-\gamma)^+}$.

Finally, $(k+2)^2 = k^2 \big((k+2)/k\big)^2 < k^2 (1 + 4/k + 4/k^2) < 4$ for $k \geqslant 4$.

So we have, for $\delta < 1$, $t > 9$ and $k > 4$:

$$\big((k+2)^2 \big(t-k\big)^{(2-\gamma)^+} + k^{-a}\big) \leqslant \big(2^{\gamma^+ - 2} t^{(2-\gamma)^+} t^\delta + t^{(2-\gamma)^+}\big).$$

In summary we know the following: given a choice of $\delta \in (0, 1)$, for a defined from δ as above, for any $t > \underline{K_a}^{2/\delta}$, then

$$\left| \int f \cdot g \circ T^t \, \mathrm{d}\mu - \int f \, \mathrm{d}\mu \int g \, \mathrm{d}\mu \right| \leqslant 2\bar{c}\, (\|f\|_{\mathcal{SP}} + \|g\|_{\mathcal{SP}})^2 \big(2^{\gamma^+ - 2} t^{(2-\gamma)^+} t^\delta + t^{(2-\gamma)^+}\big). \tag{72}$$

This is true for any such δ, with the number $R_\delta = \underline{K_a}^{2/\delta}$ depending only on δ (and on f and g). Thus, setting $C = 2\bar{c}\,(2^{\gamma^+-2}+1)$, we conclude that (72) is

$$\leqslant C(\|f\|_{\mathcal{SP}} + \|g\|_{\mathcal{SP}})^2\, t^{(2-\gamma)^{++}} \tag{73}$$

as stated. □

5. Convergence of the Ruelle operator for $\mathcal{SP}$ observables

We next prove an estimate which will provide the key to the CLT: polynomial convergence of the Ruelle operator in L^1- and L^2-norms, for $\gamma > 3$ and 4, respectively.

Theorem 5.1. *Let $\gamma > 2$; for $\bar{c}$ as above, for $f \in \mathcal{SP}$, we have*

$$\left\|\mathcal{L}_\psi^t f - \mu f\right\|_1 \leqslant \bar{c}\|f\|_{\mathcal{SP}}\, t^{(2-\gamma)^+}.$$

Proof. The argument will combine elements from the proofs of weak Bernouilli, and of the polynomial decay of correlation for $\mathcal{SP}$. Let $f_k = \sum_{P\in\mathcal{C}_k} a_P I_P$ be the k-cylinder approximation to f, as in the proof of theorem 4.1. We have

$$\left\|\mathcal{L}^t f - \mu f\right\|_1 \leqslant \left\|\mathcal{L}^t(f-f_k) - \mu(f-f_k)\right\|_1 + \left\|\mathcal{L}^t f_k - \mu f_k\right\|_1.$$

The first term here is

$$= \left\|\mathcal{L}^t(f-f_k)\right\|_1 \leqslant \left\|\mathcal{L}^t(f-f_k)\right\|_\infty \leqslant \|f-f_k\|_\infty \leqslant \mathrm{var}_k\, f \leqslant \|f\|_{\mathcal{SP}}\cdot k^{-a}$$

for all $a > 0$. For the second term,

$$\left\|\mathcal{L}^t f_k - \mu f_k\right\|_1 = \Big\|\sum_{P\in\mathcal{C}_k} a_P(\mathcal{L}^t I_P - \mu P)\Big\|_1 \tag{74}$$

$$\leqslant \|f\|_\infty \sum_{P\in\mathcal{C}_k} \left\|\mathcal{L}^t I_P - \mu P\right\|_1 \tag{75}$$

$$\leqslant \bar{c}\|f\|_{\mathcal{SP}}(k+2)^2(t-k)^{(2-\gamma)^+} \tag{76}$$

where the passage from equation (75) to (76) is exactly that of steps (68) through (69) in the proof of theorem 3.1. So

$$\left\|\mathcal{L}^t f - \mu f\right\|_1 \leqslant \|f\|_{\mathcal{SP}}\big(k^{-a} + \bar{c}(k+2)^2(t-k)^{(2-\gamma)^+}\big)$$

for all $t > k > \underline{k_a}(f)$ and for all $a > 0$.

As in the end of the proof of theorem 5.1, this implies

$$\left\|\mathcal{L}^t f - \mu f\right\|_1 \leqslant \bar{c}\|f\|_{\mathcal{SP}}\, t^{(2-\gamma)^+}$$

as we wished to show. □

For the next theorem we need a lemma, for the proof of which see the preprint version of this paper referred to earlier. This is the L^2-version of (67).

Lemma 5.2. *Let $\gamma > 2$, then for any fixed k-cylinder set P,*

$$\int \left|\mathcal{L}_\psi^t I_P(z) - \mu P\right|^2 \mathrm{d}\mu(z) \leqslant \left(\bar{c}\,\frac{\mu[w_0\ldots w_{j-1}0]}{\mu[0]}\right)^2 2(k+1)^2(t-k)^{(2-\gamma)^+}.$$

Theorem 5.3. *Let $\gamma > 2$; for $\bar{c}$ as above, for $f \in \mathcal{SP}$, we have*

$$\left\|\mathcal{L}_{\psi}^{t} f - \mu f\right\|_2 \leqslant \bar{c}\|f\|_{\mathcal{SP}}\, t^{(1-\gamma/2)^{++}}.$$

Proof. This will be similar to the proof of theorem 5.1. As there, we have two terms; the first is handled just as before. For the second, using the lemma

$$\left\|\mathcal{L}^{t} f_k - \mu f_k\right\|_2 \leqslant \|f\|_\infty \sum_{P\in\mathcal{C}_k} \left\|\mathcal{L}^{t} I_P - \mu P\right\|_2 \tag{77}$$

$$\leqslant \sqrt{2}(k+1)^2 \bar{c}(t-k)^{(1-\gamma/2)^{+}}. \tag{78}$$

The rest of the proof is just the same as for theorem 5.1. □

Corollary 5.4.

(a) Assume $\gamma > 4$; for $\bar{c}$ as above, for $f \in \mathcal{SP}$, we have

$$\sum_{t=1}^{\infty} \left\|\mathcal{L}_{\psi}^{t} f - \mu f\right\|_2 \leqslant \infty.$$

(b) If $\gamma > 3$; for $\bar{c}$ as above, for $f \in \mathcal{SP}$, we have

$$\sum_{t=1}^{\infty} \left\|\mathcal{L}_{\psi}^{t} f - \mu f\right\|_1 \leqslant \infty.$$

Also, the series

(i) $$\sum_{t=1}^{\infty} \left(\mathcal{L}_{\psi}^{t} f - \mu f\right)$$

converges in L^1 and L^2 for (a) and (b), respectively, and in either case,

(ii) $$\sum_{t=1}^{\infty} \left|\mathcal{L}_{\psi}^{t} f - \mu f\right|(x)$$

converges for μ-a.e. x.

Proof. We give the proof with assumption (b), that for (a) being similar, but using instead theorem 5.3. From theorem 5.1 we know that for any $\delta > 0$, there exists R_δ such that for all $t > R_\delta$,

$$\left\|\mathcal{L}_{\psi}^{t} f - \mu f\right\|_1 \leqslant \bar{c}\|f\|_{\mathcal{SP}}\, t^{2-\gamma+\delta}$$

we choose δ so small that $\gamma - \delta > 3$. Hence the series in (b) converges. Since

$$\left\|\sum_{t=m}^{\infty} \mathcal{L}_{\psi}^{t} f - \mu f\right\|_1 \leqslant \sum_{t=m}^{\infty} \left\|\mathcal{L}_{\psi}^{t} f - \mu f\right\|_1 < \varepsilon$$

for m large enough, the first sequence of partial sums is a Cauchy sequence and hence converges. The same argument works for $\|\cdot\|_2$.

Now (ii) follows from a simple real analysis argument: assuming (b) (which is implied by (a)), if g_i are measurable with $ct \sum \|g_i\| < C$ and $A_M = \{x : \sum |g_i(x)| > M\}$, then certainly the measure of $A_M < C/M \to 0$ as $M \to \infty$, so $\cap A_M$ has measure 0, completing the proof. □

6. The central limit theorem

We will now show how to use the above estimates to prove the central limit theorem, for functions of class $\mathcal{SP}$ and $\gamma > 3$.

We write $\mathcal{B}$ for the Borel σ-algebra of Σ^+, and define $\mathcal{B}_k = T^{-k}\mathcal{B}$. We write $\mathcal{F}$ for the collection of $\mathcal{B}$-measurable functions and $\mathcal{F}_k$ for the $\mathcal{B}_k$-measurable functions. Note that, for example, $f \in \mathcal{F}_1 \iff f(x) = f(y)$ when $T(x) = T(y)$; hence $\mathcal{F}_1$ is exactly the collection of measurable functions which are of the form $f \circ T$ with f a $\mathcal{B}$-measurable function. Defining $U\colon \mathcal{F} \to \mathcal{F}$ by $U\colon f \mapsto f \circ T$, we note that $U(\mathcal{F}_k) = \mathcal{F}_{k+1}$. Since T preserves the measure μ, for $L^2 = L^2(\Sigma^+, \mu)$, $U\colon L^2 \to L^2 \cap \mathcal{F}_1$ is an isometry.

The following is well known (see [PP90, p 27]) and is not hard to prove using the fact that $\mathcal{L}_\phi{}^*(\mu) = \mu$ (see the beginning of section 2 or [Hof77]).

Proposition 6.1. *For a normalized (measurable) potential ϕ, with eigenmeasure μ, the adjoint $U^*\colon L^2 \to L^2$ is equal to the Ruelle operator $\mathcal{L}_\phi$.*

What this means is that, when $\mathcal{L}_\phi$ is extended from the continuous functions to L^2, then it equals U^*. In other words, $\mathcal{L}_\phi$ is the L^2-dual of U.

Remark 6.1. The orthogonal projection P_k onto $L^2 \cap \mathcal{F}_k$ is often written as

$$P_k(f) = \mathbb{E}(f|\mathcal{B}_k).$$

This is the expected value of f with respect to the σ-algebra $\mathcal{B}_k$ [Bil68]. All the statements of [Via97, Liv95] are formulated in this language, which is natural and useful from the point of view of probability theory; we have chosen instead to emphasize here instead the connections with the Ruelle operator.

We point out that the next result under the hypothesis

$$\sum_{n=1}^{\infty} \|\mathcal{L}_\phi^n f\|_2 < \infty$$

was proved by Gordin [Gor69, Roz79, Bra88]. An improvement by Liverani (see [Liv95]) assumes less. We will need this weaker assumption (next theorem) for the case $\gamma > 3$ (otherwise using [Gor69] we would have only the result for $\gamma > 4$).

Theorem 6.2 (See [Via97, Liv95]). *Assume that μ is an invariant ergodic probability measure on Σ^+, with corresponding normalized potential ϕ (see before lemma 2.1), let f be a measurable function with mean 0 ($\int f \,\mathrm{d}\mu = 0$) such that*
(a) $\sum_{n=1}^{\infty} |\int (f \circ T^n) f \,\mathrm{d}\mu| < \infty$ and
(b) $\sum_{n=1}^{\infty} \|\mathcal{L}_\phi^n f\|_1 < \infty$.

Setting $u = -\sum_{n=1}^{\infty} \mathcal{L}_\phi^n f$, and defining $\tilde{f} = f - u + u \circ T$ then $\tilde{f}_k \equiv \tilde{f} \circ T^k$ is an orthogonal sequence of random variables. The variance $(\|\tilde{f}\|_2)^2$ of $\tilde{f}$ satisfies

$$\|\tilde{f}\|_2^2 = \|f\|_2^2 + 2\sum_{n=1}^{\infty} \int f \cdot (f \circ T^n) \,\mathrm{d}\mu.$$

This is 0 if and only if $f = v \circ T - v$ for some $v \in L^2(\mu)$.

We note that, in fact, theorem 1.1 in [Liv95] requires less than the above condition that $\sum_{n=1}^{\infty} \|\mathcal{L}_\phi^n f\|_1 < \infty$.

An immediate corollary of the orthogonality (see [Via97]) is that $\tilde{f}_k \equiv \tilde{f} \circ T^k$ is a (reversed) martingale difference sequence with finite variance. The central limit theorem is known for such processes. This is due independently [Bil68] to Billingsley and Ibragimov [Bil61, Ibr63]; see also [DG86]. The functional version is true as well as the standard CLT [Bil68].

Theorem 6.3 (Central limit theorem for martingales; Billingsley and Ibragimov). *Let X_i be a stationary, ergodic stochastic process such that the partial sums S_t form a martingale. Then the central limit theorem (and functional CLT) hold for X_i with variance $\sigma^2 = \text{variance}(X_i)$.*

This result is then transferred from $\tilde{f}$ to f via the cohomology equation (see [Liv95]), and we have:

Theorem 6.4. *Under the assumptions of theorem 6.2, the central limit theorem and functional CLT hold for the sequence $f_k \equiv f \circ T^k$.*

Theorem 6.5. *For $\gamma > 3$, the central limit theorem and functional CLT hold for mean-zero observables f in $\mathcal{SP}$; the variance is*

$$\sigma^2 = \int f^2 \, \mathrm{d}\mu + 2 \sum_{n=1}^{\infty} \int f \cdot (f \circ T^n) \, \mathrm{d}\mu$$

and this is 0 if and only if $f = v \circ T - v$ for some $v \in L^2(\mu)$.

Proof. Theorems 4.1 and corollary 5.4(b) give us conditions (a) and (b) of theorem 6.2. Hence combining that result with theorem 6.4 finishes the proof. □

Remark 6.2. It is instructive to see a direct proof of the key hypotheses (a) and (b) of theorem 6.2 for the specific case of $f = I_{[1]} - \mu[1]$, using only estimate (20), that for all $z \in M_s$,

$$|\mathcal{L}_{\psi}^n f(z)| \leqslant \hat{c}_\gamma \big((s+n)^{2-\gamma} + B(s,n) \big) \tag{79}$$

where

$$B(0,n) = 0 \qquad \text{and} \qquad B(s,n) = \left(\frac{s+n}{s} \right)^{1-\gamma} \qquad \text{for} \quad s \geqslant 1.$$

To prove (a), note that

$$\left| \int (f \circ T^n) \, f \, \mathrm{d}\mu \right| = \left| \int f \, \mathcal{L}_\phi^n f \, \mathrm{d}\mu \right| \leqslant \int |f \, \mathcal{L}_\phi^n f| \, \mathrm{d}\mu \leqslant \int |\mathcal{L}_\phi^n f| \, \mathrm{d}\mu.$$

And we have, since $\mu(M_s) = \mu_s \leqslant s^{1-\gamma}$ while also $\sum \mu_s = 1$, that

$$\begin{aligned}
\frac{1}{\hat{c}_\gamma} \int |\mathcal{L}_\phi^n f| \, \mathrm{d}\mu &\leqslant \sum_{s=0}^{\infty} \mu_s \big((s+n)^{2-\gamma} + B(s,n) \big) \\
&\leqslant \sum_{s=0}^{\infty} \mu_s (s+n)^{2-\gamma} + \sum_{s=1}^{\infty} \left(\frac{s+n}{s} \right)^{1-\gamma} s^{1-\gamma} \\
&\leqslant \sum_{s=0}^{\infty} \mu_s n^{2-\gamma} + \sum_{s=1}^{\infty} (s+n)^{1-\gamma} \leqslant 2n^{2-\gamma}.
\end{aligned}$$

To prove (b), we have that

$$\begin{aligned}
\frac{1}{\hat{c}_\gamma} \sum_{n=1}^{\infty} \|\mathcal{L}_\phi^n f\|_1 &\leqslant \sum_{n=1}^{\infty} \sum_{s=0}^{\infty} (s+n)^{2-\gamma} \, s^{1-\gamma} + \sum_{n=1}^{\infty} \sum_{s=1}^{\infty} \left(\frac{s+n}{s} \right)^{1-\gamma} s^{1-\gamma} \\
&\leqslant \sum_{s=0}^{\infty} s^{1-\gamma} (s+1)^{3-\gamma} + \sum_{s=1}^{\infty} \sum_{n=1}^{\infty} (s+n)^{1-\gamma} \\
&\leqslant \sum_{s=0}^{\infty} s^{4-2\gamma} + \sum_{s=1}^{\infty} (s+1)^{2-\gamma} < \infty.
\end{aligned}$$

Note added in proof. After this paper was submitted we were informed by the referees about papers with results of a similar nature [Iso99, Huy, Che95]. A long version of the present paper with full details can be found in http://ime.usp.br/∼afisher.

Acknowledgments

AF was supported by CNPQ and FAPERGS; AL was partially supported by PRONEX-Sistemas Dinâmicos (CNPQ). We thank F Ledrappier for pointing out to us an error in an early version of this paper. We also wish to thank M Sebastiani for helpful conversations related to this paper. We would like to give special thanks to E Brietzke for supplying us with a result [Bri00] that was just what we needed when proving the lower bound estimate of our example.

Appendix. An integral estimate

Let $\alpha, \beta < 0$ and define

$$c_n = 1^\alpha (n-1)^\beta + 2^\alpha (n-2)^\beta + \cdots + (n-1)^\alpha 1^\beta = \sum_{k=1}^{n-1} k^\alpha (n-k)^\beta = \sum_{k=1}^{n-1} k^\beta (n-k)^\alpha. \tag{A1}$$

We set

$$f(x) = x^\alpha (1-x)^\beta$$

and write $\int f$ for $\int_0^1 f(x)\,\mathrm{d}x$. Note that this function is integrable exactly for $\alpha, \beta \in (-1, 0)$.

Lemma A1. *Suppose $\alpha, \beta \in (-\infty, -1)$. Then*

$$c_n \approx \left(\int f\right) n^{\alpha+\beta+1}$$

i.e. for $\check{c}_{\alpha,\beta} \equiv \int f$, and any constants $\check{c}^+_{\alpha,\beta}, \check{c}^-_{\alpha,\beta}$ with $\check{c}^-_{\alpha,\beta} < \check{c}_{\alpha,\beta} < \check{c}^+_{\alpha,\beta}$,

$$\check{c}^-_{\alpha,\beta} \cdot n^{\alpha+\beta+1} \leqslant \sum_{k=1}^{n-1} k^\alpha (n-k)^\beta = \sum_{k=1}^{n-1} k^\beta (n-k)^\alpha \leqslant \check{c}^+_{\alpha,\beta} \cdot n^{\alpha+\beta+1}$$

for n large enough.

When the function f is not integrable, for an upper estimate of c_n, the first and last terms in that sum dominate; these are $(n-1)^\beta$ and $(n-1)^\alpha$. We see this by a truncated lower Riemann sum estimate, as follows.

Lemma A2. *Suppose $\alpha < \beta < 0$ and that $\alpha < -1$. Then there exists a positive constant $\bar{c}_{\alpha,\beta}$ such that for all n:*

$$c_n = \sum_{k=1}^{n-1} k^\alpha (n-k)^\beta = \sum_{k=1}^{n-1} k^\beta (n-k)^\alpha \leqslant \bar{c}_{\alpha,\beta} (n-1)^\beta.$$

The proof of the next statement is similar, except simpler because we only need to deal with the singularity at 0.

Lemma A3. *Let $\alpha < -1$ and $\alpha < \beta < 0$ and assume $k+1 \leqslant t \cdot \lambda = t \cdot \alpha/(\alpha+\beta)$. Then*

$$\sum_{j=2}^{k+1} j^\alpha (t-j)^\beta \leqslant \bar{c}_{\alpha,\beta} t^\beta.$$

Note. For $\alpha = (1-\gamma)$ and $\beta = (2-\gamma)$, for $\gamma > 2$, $\lambda = (1-\gamma)/(3-2\gamma) < 2$ and the condition $k+1 < \lambda t$ becomes $t > (k+1)/2$. When we apply this lemma, this is satisfied since we have, in fact, $t \geqslant k+1 > (k+1)/2$.

Lemma A4. *Let $\alpha < \beta < 0$ and assume $\alpha < -1$. Then $\exists\, \hat{c}_{\alpha,\beta} > 0$ such that for all $s, t \geqslant 1$, we have*

$$\sum_{j=1}^{t-1} (s+j)^\alpha (t-j)^\beta$$

respectively

$$\begin{aligned} &\leqslant \hat{c}_{\alpha,\beta}\big(s^{\alpha+1}(s+t)^\beta + (s+t)^\alpha\big) && \text{for} \quad \beta < -1 \\ &\leqslant \hat{c}_{\alpha,\beta}\big(s^{\alpha+1}(s+t)^\beta + (s+t)^\alpha \log(s+t)\big) && \text{for} \quad \beta = -1 \\ &\leqslant \hat{c}_{\alpha,\beta}\big(s^{\alpha+1}(s+t)^\beta + (s+t)^{1+\alpha+\beta}\big) && \text{for} \quad \beta \in (-1,0). \end{aligned}$$

Definition of constants. We use the following notation in the paper:

$$\begin{aligned} &\bar{c}_1 = \bar{c}_{\alpha,\beta} && \text{for} \quad \alpha = (1-\gamma) \quad && \beta = (2-\gamma) \\ &\bar{c}_2 = \bar{c}_{\alpha,\beta} && \text{for} \quad \alpha = -\gamma && \beta = (2-\gamma) \\ &\bar{c}_3 = \hat{c}_{\alpha,\beta} && \text{for} \quad \alpha = -\gamma && \beta = (2-\gamma). \end{aligned}$$

All the other constants which appear in the paper are derived from these basic constants.

References

[Aar97] Aaronson J 1997 *An Introduction to Infinite Ergodic Theory (Mathematics Surveys and Monographs)* (AMS)

[ADF92] Aaronson J, Denker M and Fisher A M 1992 Second order ergodic theorems for ergodic transformations of infinite measure spaces *Proc. AMS* **114** 115–28

[ADU93] Aaronson J, Denker M and Urbanski M 1993 Ergodic theory for Markov fibered systems and parabolic rational maps *Trans. AMS* **337** 495–548

[AF] Arnoux P and Fisher A M 2001 The Ruelle–Perron–Frobenius theorem for random and nonstationary subshifts in preparation

[Bar] Bary N K A treatise in trigonometric series, vol II

[BFG99a] Bressaud X, Fernandez R and Galves A 1999 Decay of correlations for non-Hölderian dynamics. A coupling approach *Electron J. Prob.* **4** 1–19

[BFG99b] Bressaud X, Fernandez R and Galves A 1999 Speed of d-convergence for Markov approximations of chains with complete connections. A coupling approach *Stoch. Process. Appl.* **83** 127–38

[BG95] Bogenschütz T and Gundlach M 1995 Ruelle's transfer operator for random subshifts of finite type *Ergod. Theor. Dynam. Syst.* **15** 413–47

[Bil61] Billingsley P 1961 The Lindeberg–Lévy theorem for martingales *Proc. AMS* **12** 788–92

[Bil68] Billingsley P 1968 *Convergence of Probability Measures* (New York: Wiley)
[Bow75] Bowen R 1975 *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms (Lecture Notes in Mathematics 470)* (Berlin: Springer)
[Bra88] Bradley R C On some results of M I Gordin: a clarification of a misunderstanding *J. Theor. Prob.* **1** 1988
[Bre68] Breiman L 1968 *Probability* (Reading, MA: Addison-Wesley)
[Bri00] Brietzke E 2000 Monotonicity of coefficients of reciprocal power series *Real Analysis Exchange* to be published
[Bro94] Broise A 1994 Aspects stochastiques de certains systems dynamiques: transformations dilatantes de l'intervale, fractions continues multidimensionelles *Thesis* University de Rennes
[Bro96] Broise A 1996 Fractions continues multidimensionelles et lois stables *Bull. Soc. Math. France* **124** 97–139
[CF90] Collet P and Ferrero P 1990 Some limit ratio theorem related to a real endomorphism in the case of a neutral fixed point *Ann. Inst. H Poincaré* **52** 283–301
[CG93] Collet P and Galves A 1993 Statistics of close visits to the indifferent fixed point of an interval map *J. Stat. Phys.* **52** 459–78
[CGS92] Collet P, Galves A and Schmitt B 1992 Unpredictability of the ocurrence of a laminar period in a model of temporal intermitency *Ann. Inst. H Poincaré* **57** 319–31
[Che95] Chernov N 1995 Limit theorems ands markov approximations for chaotic dynamical systems *Prob. Theor. Relat. Fields* **101** 321–62
[DG86] Durr D and Goldstein S 1986 *Remarks on The CLT for Weakly Dependent Random Variables (Lecture Notes in Mathematics 1158)* (Berlin: Springer) pp 104–18
[Fel49] Feller W 1949 Fluctuation theory of recurrent events *Trans. AMS* **67** 98–119
[Fel71] Feller W 1971 *Probability Theory* vol 2 (New York: Wiley)
[FLa] Fisher A M and Lopes A 2000 Invariance principles in log density and convergence to stable flows *USP preprint*
[FLb] Fisher A M and Lopes A 2000 Self-similar return sets and stable laws for some maps with an indifferent fixed point *USP preprint*
[FS79] Ferrero P and Schmitt B 1979 Ruelle's Perron–Frobenius theorem and projective metrics *Coll. Math. Soc. János Bolyai (Random Fields)* **27** 333–6
[FS88] Ferrero P and Schmitt B 1988 Produits aléatoires d'operateurs matrices de transfert *Prob. Theor. Relat. Fields* **79** 227–48
[Gor69] Gordin M I 1969 The central limit theorem for stationary processes *Dokl. Akad. Nauk SSSR* **188** 1174–6
[GS92] Grimett G and Stirzaker D 1992 *Probability and Random Processes* (Oxford: Oxford University Press)
[Har49] Hardy G 1949 *Divergent Series* (Oxford: Oxford University Press)
[Hof77] Hofbauer F 1977 Examples for the non-uniqueness of the equilibrium states *Trans. AMS* **228** 133–41
[Huy] Huyi H Decay of correlations for piecewise smooth maps with indifferent fixed points *Preprint* http://www.math.psu.edu/hu/papers.html
[Ibr63] Ibragimov I A 1963 A central limit theorem for a class of dependent random variables *Theor. Prob. Appl.* **8** 83–9
[Iso99] Isola S 1999 Renewal sequences and intermitency *J. Stat. Phys.* **97** 263–80
[Kal28] Kaluza T 1928 *Matematishe Zeitschrift* vol 28, pp 161–1170
[Kea72] Keane M 1972 Strongly mixing g-measures *Invent. Math.* **16** 309–24
[KMS97] Kondah, Maume and Schmitt B 1997 Vitesse de convergence vers l'etat d'equilibre pour le dynamique Markovienn non Hölderiennes *Ann. Inst. H Poincaré* **33** 675–95
[KT75] Karlin S and Taylor R 1975 *A First Course in Stochastic Processes* (New York: Academic)
[Led74] Ledrappier F 1974 Principe variationnel et systemes dynamiques symboliques *Z. Wahr.* **30** 185–202
[Liv95] Liverani L 1995 Central limit theorem for deterministic systems *Ann. Math.* **142** 239–301
[Lop90] Lopes A 1990 A first order level-2 phase transition in thermodynamic formalism *J. Stat. Phys.* **60** 395–411 (see appendix)
[Lop93] Lopes A 1993 The Zeta function, non-differentiabilty of pressure and the critical exponent of transition *Adv. Math.* **101** 133–65
[LSV93] Lambert A, Siboni S and Vaienti S 1993 Statistical properties of a non-uniformly hyperbolic map of the interval *J. Stat. Phys.* **71** N 1305–30
[LSV99] Liverani C, Saussol B and Vaienti S 1999 A probabilitic aproach to intermitency *Ergod. Theor. Dynam. Syst.* **11** 671–86
[Man80] Manneville P 1980 Intermittency, self-similarity and $1/f$ spectrum in dissipative dynamical systems *J. Physique* **41** 1235–43
[Mor93] Mori M 1993 On the intermittency of a piecewise linear map (takahashi model) *Tokyo J. Math.* **16** 411–28
[Pol] Pollicott M 2000 Rate of mixing for potentials of summable variation *Trans. Am. Math. Soc.* **352** 843–53

[Pop79] Popov N N 1978–79 On the rate of convergence for countable Markov chains *SIAM, Theor. Prob. Appl.* **24** 401–5

[Pos79] Postnikov A 1979 Tauberian theory and its applications *Proc. Steklov Inst. Math.*

[PP90] Parry W and Pollicott M 1990 Zeta functions and the periodic orbit structure of hyperbolic dynamics *Asterisque* **187–8**

[PS92] Prellberg and Slawny 1992 Maps of the interval with indifferent points: thermodynamic formalism and phase transitions *J. Stat. Phys.* **66** 503–14

[PY] Pollicott M and Yuri M 2001 Statistical properties of maps with indifferent periodic points *Commun. Math. Phys.* **217** 503–20

[Roz79] Rozanov D 1979 *Stochastic Processes* (Berlin: Springer)

[Via97] Viana M 1997 Stochastic behaviour of deterministic attractors *Col. Brasileiro de Matemática (IMPA)*

[Wal75] Walters P 1975 Ruelle's operator theorem and g-measures *Trans. AMS* **214** 375–87

[You99] Young L-S 1999 Recurrence times and rates of mixing *Israel J. Math.* **110** 153–88

[Yur97] Yuri M 1997 On the convergence to equilibrium states for certain non-hyperbolic systems *Ergod. Theor. Dynam. Syst.* **17** 977–1000